

5G and Edge AI for Smart Industries: Technologies Enabling Real-Time Industrial Intelligence

Karthik Kumar Vaigandla^{1,*}

¹ Associate Professor, Electronics and Communication Engineering, Balaji Institute of Technology and Science, Warangal, Telangana, India - 506331

Corresponding Author: Karthik Kumar Vaigandla. Email: ykvaigandla@gmail.com

<https://doi.org/10.69996/jsihs.2026010>

Abstract: The merging of 5G communication and Edge Artificial Intelligence (Edge AI) is driving smart industrial systems toward a revolutionary change in how they handle real-time data processing, achieve extremely fast communication speeds, and make intelligent decisions directly at the network edge. This analysis examines how 5G and Edge AI work together to power cutting-edge industrial applications like predictive maintenance, autonomous robotics, smart manufacturing, and industrial automation systems. The combination of 5G's ultra-reliable low-latency communication (URLLC), enhanced mobile broadband (eMBB), and massive machine-type communications (mMTC) features, paired with Edge AI's distributed intelligence and quick response capabilities, creates opportunities for better operational performance, increased flexibility, and stronger system reliability. This paper presents a thorough examination of key enabling technologies such as network slicing, edge computing frameworks, AI models designed for edge environments, and integration systems. The article also addresses major implementation obstacles including security concerns, scalability requirements, interoperability issues, and energy consumption challenges. Additionally, it explores current research developments and future pathways for creating completely intelligent and connected industrial environments.

Keywords: 5G, Edge AI, Smart Industries, Internet of Things, Communication

1. Introduction

Recent years have witnessed remarkable progress in wireless communication and computational technologies, fundamentally reshaping traditional industries into sophisticated, intelligent ecosystems. Two primary drivers leading this change are Fifth Generation (5G) mobile communication networks and edge computing systems. When these technologies combine within industrial settings, they enable the development of smart industries that prioritize real-time monitoring, intelligent decision-making, and autonomous control to deliver process automation, operational efficiency, and enhanced system reliability. 5G networks deliver capabilities that surpass all previous cellular generations, particularly through ultra-reliable low-latency communications (URLLC), enhanced mobile broadband (eMBB), and massive machine-type communications (mMTC) [1–3]. These characteristics prove vital for satisfying the demanding requirements of time-critical industrial applications, such as closed-loop control systems, remote maintenance operations, autonomous guided vehicles (AGVs), and digital twin implementations. Beyond these core features, 5G technologies including end-to-end quality of service (QoS), network slicing, and deterministic latency deliver exceptional communication reliability and adaptability establishing their relevance across diverse industrial domains, spanning discrete manufacturing through process industries.



Edge computing acts as an essential partner to 5G technology, moving computational power and data storage closer to where data originates usually at sensors, actuators, and industrial control devices. When data gets processed locally, edge computing lightens the load on centralized data centers, cuts down transmission delays, and improves how quickly industrial control systems respond. This becomes especially important in industrial settings where applications require high-quality data collection, instant analytics, anomaly identification, and rapid response mechanisms. When 5G and edge computing work together, they create highly distributed and responsive industrial network structures that can handle mission-critical operations. Take a smart factory as an example a robotic arm can carry out task instructions with response times measured in milliseconds, while nearby edge nodes analyze vibration data instantly to forecast possible equipment breakdowns. In the same way, within logistics operations, self-driving vehicles can navigate changing environments by using locally analyzed sensor information while staying connected to central management systems through 5G networks. [4] provides a comprehensive overview of Multi-access Edge Computing (MEC), including its fundamental concepts, enabling technologies, applications, and associated challenges. The study emphasizes the significance of MEC in enabling low-latency, real-time, and intelligent services for next-generation communication networks. It also discusses key research challenges such as resource management, security, scalability, mobility support, and interoperability, while highlighting the increasing role of MEC in emerging technologies such as 5G, IoT, artificial intelligence, and smart industrial systems.

The review paper being examined concentrates on how 5G and edge computing come together in intelligent industrial systems. It explores sub-network designs and system frameworks, alongside supporting technologies like network slicing, edge coordination, and predictable networking, emphasizing their contributions to industrial automation. The paper also covers integration approaches and system-wide obstacles such as response time requirements, expandability, and protection concerns while pinpointing current developments and future research paths. The paper delivers a thorough examination of how these advancing technologies enable real-time industrial intelligence across various industries.

2. 5G Communication

The 5G of the mobile communication network technology is designed to provide augmented information rates, ultra-reduced latency, high spectral effectiveness and gigantic vehicular connection. It draws on the 4G LTE capabilities and introduces radical changes in radio access technology, network architecture and spectrum efficiency [6]. The 5G communication plans involve a set of hi-tech technologies, which considerably boost the potentiality, capacity, and flexibility of the wireless networks. Of all the technologies, one of the most important is 5G New Radio (NR) that is the international standard that defines the 5G air interface. Scalable numerology provisioned in 5G NR where the subcarrier spacing $15 \text{ kHz} * 2^n$ ($n = 0, 1, 2, 3, \dots$) offers the capabilities to support a diverse spectrum of deployment scenarios, including low-latency and high-throughput applications. It is also going to employ superior modulation systems like 256-QAM and implement both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD) to help manage spectrum usage. Figure 1 shows key enabling technologies for 5G.

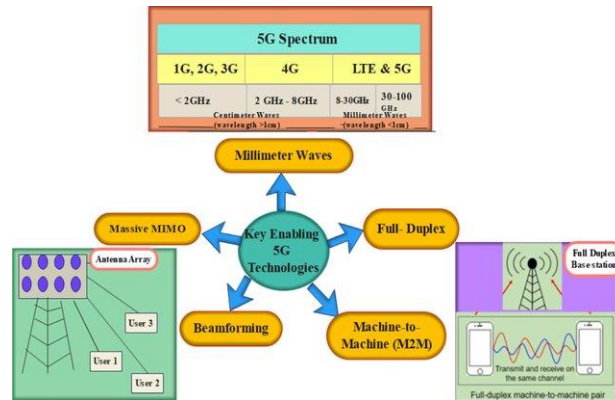


Figure 1: Key enabling technologies for 5G [7]

Massive Multiple Input Multiple Output (mMIMO), which deploys large numbers of antennas at the base station, and serves multiple users simultaneously [8]. This improves spectral efficiency and improves system capacity using spatial multiplexing. Massive MIMO is also coupled with digital beamforming making the beam to the user himself. This enhances signal quality and eliminates the interference. As well as in millimeter-wave (mmWave) communications, where frequencies (usually above 24 GHz) are used to achieve large bandwidths to support ultra-high data rates, beamforming is also critical. Nevertheless, mmWave signal has a high path loss and a weak penetration, and has to be deployed densely in small cells and operate advanced signal-focusing antenna arrays. Small Cells and Heterogeneous Network (HetNs), to alleviate the reduced coverage of mmWave and increase a network capacity, 5G is based on small cell hyperdensities and combining of different types of cells, creating HetNets. The macro cells, micro cells and femtocell work together to provide seamless coverage. The other enabling technology is network slicing and this enables us to split the physical network into several logical slices. Individual slices are also customized to address certain service needs including reduced latency and increased reliability. This is achieved by Software-Defined Networking (SDN) and Network Function Virtualization (NFV), thus creating a decoupling of hardware and software so as to allow programmable and flexible control of the network. Besides, edge computing forms part of the 5G architecture in order to decentralize computation to the user so as to reduce latency and in turn backhaul.

The core technologies that make 5G possible include scalable Orthogonal Frequency Division Multiplexing (OFDM)-based New Radio (NR), massive Multiple-Input Multiple-Output (mMIMO), millimeter-wave (mmWave) communication, beamforming, network slicing, and virtualization. These technologies work together to enable various applications like enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive Internet of Things (mIoT) connectivity [9–10]. Each technology brings major improvements compared to earlier wireless generations, delivering higher data speeds, better responsiveness, and improved energy performance. 5G networks operate using three primary frequency ranges low, mid, and high bands which correspond to frequencies under 1 GHz, between 1–6 GHz, and above 24 GHz. The highest frequency range is commonly called the mmWave band. Low-band frequencies provide broad coverage with excellent signal penetration through buildings and obstacles, making them perfect for rural areas and wide-scale deployments. However, these frequencies deliver slower data speeds. Mid-band frequencies create a good balance between coverage area and data capacity. Common mid-band frequencies like 3.5 GHz and 4.9 GHz help deliver improved mobile broadband services in city environments. High-band frequencies

(mmWave), spanning from 24 GHz to 100 GHz, offer extremely fast data rates and enormous bandwidth capacity [11–12]. The downside is their short transmission range and weak signal strength, which means more base stations must be installed close together to maintain consistent coverage and performance.

The 5G Core Network (5GC) features a Service-Based Architecture (SBA), with improved flexibility, scalability, and cloud-native compared to the 4G Evolved Packet Core (EPC). Among these are the Access and Mobility Management Function (AMF), controlling registration of the user and management of the mobility, the Session Management Function (SMF) to control and manage session setup, and the User Plane Function (UPF) to control the forwarding of user data [13]. These functions are decoupled in 5GC to allow network slicing and heterogeneous access technology support. Network slicing is the use of a shared physical network to build multiple logical networks with different service optimizations, e.g. ultra-reliable communication or high bandwidth. The architecture of the 5G Core Network is shown in figure 2.

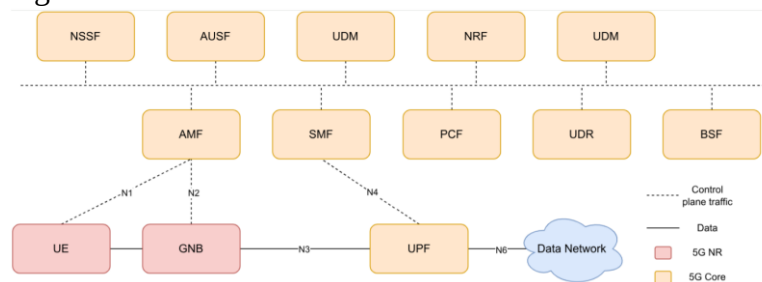


Figure 2: Architecture of the 5G Core Network [14]

One of the characteristic features of 5G networks is latency and data rate performances. The 5G theoretical maximum data rates are 20 Gbps downstream and 10 Gbps upstream, but in practice, will be significantly lower. In the case of latency, 5G targets end-to-end delay below 1 millisecond that is essential in areas of application such as remote surgery, industrial automation, and autonomous driving. All these advances in latency and throughput make possible new categories of services, including eMBB with higher-resolution video streaming and immersive AR/VR applications, URLLC suitable to mission-critical applications, and mMTC scale up to thousands of connected devices per square kilometer. The transformative role of 5G technology in accelerating intelligent automation and industrial digitization. The study highlights how the high-speed connectivity, ultra-low latency, massive device connectivity, and enhanced reliability of 5G networks support advanced industrial applications such as smart manufacturing, Industrial IoT, autonomous systems, predictive maintenance, and real-time data analytics. It also emphasizes the integration of emerging technologies including artificial intelligence, cloud computing, robotics, and edge computing in enabling Industry 4.0 environments and next-generation smart industries [15]. The features of 5G mobile technology is represented in figure 3.

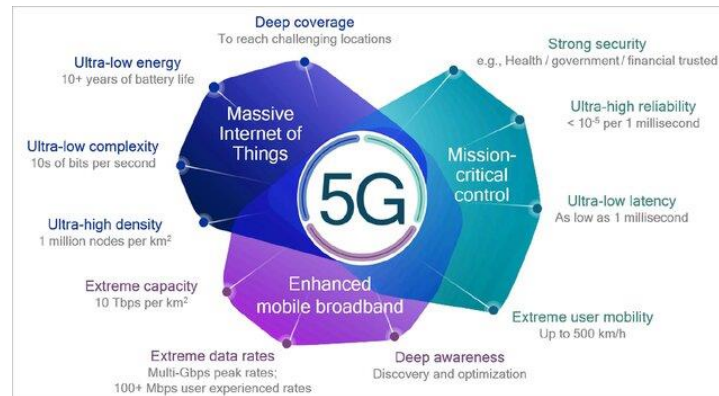


Figure 3: Features of 5G mobile technology [16]

By 2030, 5G technology will move to become more closely aligned with non terrestrial connectivity low Earth orbit (LEO) satellite systems to support ubiquitous coverage. New technologies like integrated access and backhaul (IAB), terahertz communication and reconfigurable intelligent surfaces (RIS) are under development to improve coverage, energy efficiency and spectrum utilization. Additionally, the development of virtualization and orchestration will also allow dynamic network slicing, edge computing and better quality of service with time-sensitive applications. Such trends make 5G the foundation of the next generation connectivity in industrial automation, smart cities, healthcare as well as in intelligent transportation systems.

5G technology is designed to support Ultra-Reliable Low-Latency Communication (URLLC), with theoretical latency as low as 1 ms under ideal conditions. However, in practical industrial deployments, the actual latency may vary depending on several factors such as network congestion, distance from the base station, edge computing availability, interference, spectrum allocation, and hardware limitations. Many research articles incorrectly present the 1 ms latency as consistently achievable in all scenarios, which leads to inconsistent or misleading claims. In real-world industrial environments, end-to-end latency is often higher, typically ranging from a few milliseconds to tens of milliseconds depending on the application and network architecture. Therefore, manuscripts discussing 5G performance should clearly distinguish between theoretical latency targets and practical deployment performance. Including this clarification improves the technical accuracy and credibility of the study.

3. Edge AI

Artificial Intelligence (AI) refers to the capability of machines or computer systems to emulate human cognitive functions such as learning, reasoning, problem-solving, pattern recognition, and decision-making. AI systems are generally classified into Narrow AI, which is designed for specific task-oriented applications, and General AI, which theoretically possesses human-like intelligence capable of performing a wide range of intellectual tasks. In modern computational environments, AI enables systems to process large volumes of data, extract meaningful insights, and make autonomous decisions with minimal human intervention. Owing to these capabilities, AI has found extensive applications in domains such as computer vision, speech recognition, natural language processing (NLP), robotics, healthcare, industrial automation, and intelligent communication systems [17].

Edge AI refers to the deployment and execution of artificial intelligence (AI) algorithms directly on edge devices or edge servers located close to the data source, rather than relying

entirely on centralized cloud computing platforms. In Edge AI systems, data generated by sensors, industrial machines, cameras, or IoT devices is processed locally to enable real-time decision-making with minimal communication delay. This approach reduces latency, lowers bandwidth consumption, enhances data privacy, and improves system reliability, especially in time-sensitive industrial applications. Edge AI combines edge computing infrastructure with machine learning and deep learning techniques to support intelligent operations such as predictive maintenance, anomaly detection, autonomous control, and real-time monitoring. Unlike conventional cloud-based AI systems, Edge AI minimizes the need for continuous data transmission to remote servers, making it more suitable for industrial environments where fast response, secure communication, and uninterrupted operation are essential. The integration of Edge AI with advanced communication technologies such as 5G further enhances the capability of industrial systems by enabling efficient distributed intelligence and low-latency connectivity. Edge AI is a specialized branch of AI in which machine learning and inference capabilities are deployed directly on edge devices, including IoT nodes, embedded systems, smartphones, sensors, autonomous vehicles, and industrial controllers. Unlike conventional cloud-centric AI architectures, where data must be transmitted to remote servers for processing, Edge AI performs computation and decision-making locally at or near the data source. This decentralized processing architecture minimizes communication delays, reduces bandwidth consumption, enhances data privacy and security, and enables real-time intelligent operations even in environments with limited or intermittent network connectivity [18]. Therefore, Edge AI combines the intelligence of AI algorithms with the low-latency and distributed nature of edge computing, making it highly suitable for time-sensitive applications such as smart healthcare, autonomous systems, industrial IoT, and next-generation wireless communication networks. An overview of the Edge AI architecture is illustrated in Figure 4.

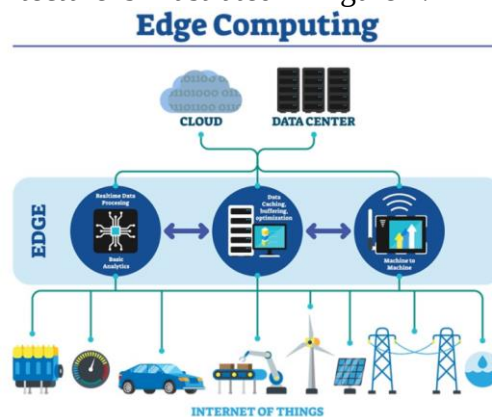


Figure 4: An Overview of Edge AI.

The most important elements of Edge AI systems are the edge devices, i.e., the devices that have computational capability (e.g., microcontrollers, DSPs, FPGAs or dedicated AI accelerator), embedded storage, communication systems, and firmware designed to run local inference. Typical AI models are trained in a cloud or on GPU or TPU-supported infrastructures and then become compressed or quantized to execute on edge devices with constrained computational resources. These models are implemented by means of optimized run-time systems like TensorFlow Hardware Accelerators or ONNX Runtime. Edge AI employs numerous AI algorithms depending on the application in use. Some of the more common algorithms are decision tree, support-vector machine, convolutional neural network (CNN) on images, a recurrent neural network (RNN) and long-short term memory (LSTM) on time-series

data, and clustering applications such as the k-means algorithm on unsupervised data. To allow efficient deployment to edges, model optimization strategies like pruning, quantization and distillation of knowledge are needed to minimize the computational load and memory requirements [19]. Edge AI has a lot of advantages. The first one is that it has low latency because inference occurs locally without round-trip traffic to the cloud. This is essential in such applications as autonomous cars, factory automation/robotics, and medical monitoring, where every millisecond matters. The efficiency of bandwidth is a second benefit of the technique, since the raw data does not have to be continually forwarded to central servers. Privacy and security are also improved since the sensitive data is stored in the device, and there is minimal exposure to possible attacks during data transmission. Edge AI can also be used to provide more resilience and availability, so that similar to disconnected or low-connectivity scenarios, the devices can keep operating.

Nevertheless, a few issues become obstacles to the utilization of Edge AI. This consists of less processing capability and storage in edge devices hence complexity in deployment of models with great complexity is cumbersome. The concern regarding energy efficiency is especially about battery-powered equipment's. Model accuracy is a tradeoff that is not trivial when compressing models to execute on the edge. Besides, standardization and model-portability is hindered by the diversity of the hardware platforms. The issue is also that security at the edge has inadequacies with physical access and the limited multiple layers as in cloud environments. Future Edge AI is ready to explode and hardware is set to provide the ability to nurture more capable and efficient inference engines. More advanced and autonomous edge systems will be made possible with the rising trend in neuromorphic computing, low-power AI-processing, and integrated sensing-processing hardware. The development of common Edge AI framework interoperability, better tooling of edge deployments, and the integration of edge computing with 5G and beyond 5G networks will enable novel possibilities in industrial automation, intelligent transport systems, smart cities, and connected healthcare. The convergence of edge intelligence and real-time analytics will transform the nature of processing and using data at its source and allow highly responsive, adaptive systems across very diverse areas.

Edge AI offers several advantages in industrial environments by enabling data processing directly at or near the source of data generation rather than relying entirely on centralized cloud servers. Recent studies have shown that Edge AI significantly reduces communication latency and bandwidth usage by processing critical information locally, which is particularly beneficial for real-time industrial applications such as predictive maintenance, robotic control, quality inspection, and autonomous monitoring systems. Experimental results reported in the literature indicate that edge-based processing can reduce response time by several milliseconds compared to cloud-based architectures while also improving system reliability during network interruptions. In addition, Edge AI enhances data privacy and security because sensitive industrial data can be analyzed locally without continuous transmission to external servers. Researchers have also demonstrated that integrating Edge AI with 5G networks improves operational efficiency through faster decision-making, optimized resource utilization, and reduced energy consumption. These advantages make Edge AI a promising solution for supporting scalable, intelligent, and low-latency industrial automation systems.

4. Enabling Technologies

4.1 5G Communication Capabilities

The 5G wireless mobile communication system brings an unprecedented change in wireless communication, providing a viable structure to industrial applications, which require large data rates, a low transmission delay and large connection, and network performance and customization. The overall architecture of the 5G is oriented towards three major categories of services namely improved eMBB, URLLC, and mMTC [20]. Such capabilities are also enhanced by the architectural advancement like network slicing which enables customized network performance to support various industrial use cases. Figure 5 shows the capabilities and applications of 5G.

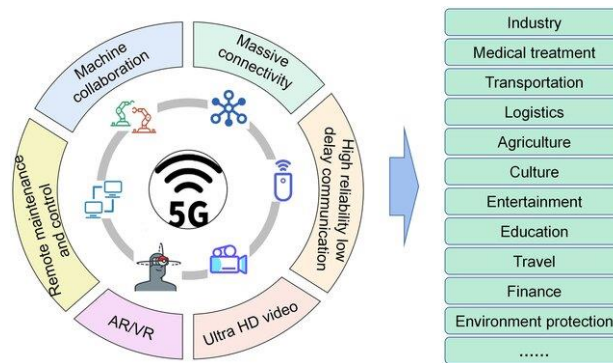


Figure 5: Capabilities and applications of 5G [21]

eMBB: eMBB is characterized by high data rates and capacity to transmit large amount of data at high rates. Within industrial settings, eMBB facilitates applications that need high bandwidth like high-resolution video surveillance, AR-assisted maintenance and digital twin inspection. Having the peak Gbps and the user-perceived data rates up to several hundred Mbps, eMBB allows smooth interactions among industrial devices, remote control terminals, and cloud assets. Use of MIMO antenna systems and beamforming as well as the mmWave frequency bands mean that the high throughput is supported even on congested factory floors or containing metallic objects.

URLLC: URLLC is designed as a low latency (latency of 1 millisecond or less) low packet error rate (less than 10^{-5} packet error rate) end to end solution. It is essential in time-sensitive industrial control e.g., robotic control loops, machine-tool motion control and safety-related signaling. 5G uses short transmission time intervals (TTIs), rapid uplink access, and preemptive scheduling as features in order to achieve such reliability and low latency. URLLC is especially useful in applications, which need deterministic operation, like synchronized multi-robot coordination, real-time feedback, remote actuation of closed-loop industrial control systems.

mMTC: mMTC concerns itself in facilitating connectivity in large volumes of low-cost, low-Power devices that are distributed in industrial ecosystems. Such devices encompass sensors, actuators, meters and tracking gadgets that as a rule permanently track parameters, similar to temperature, vibration, humidity or pressure. It is compatible with extended battery life and extensive indoor coverage that one requires in extensive applications in industrial complexes, warehouses or mines.

Network Slicing: A key capability on 5G architecture is network slicing that enables multiple virtualized and isolated logical networks to be realized over a common physical infrastructure. The performances of particular industrial applications can be made to fit in each slice as demanded. As an example, one slice may be set up with a large bandwidth, and moderate latency to perform video surveillance, and another set up with tight latency and reliability

requirements to manage a machine. Network slicing also facilitates the optimization of resources and flexibility of operations and will see enterprises use a private 5G network to suit varied operations without jeopardizing security or performance.

Aiding technologies and Protocol Improvements: These 5G features are enabled by an advanced radio access technologies, flexible numerologies, and spectrum adaptability to exploit all their potential. Adaptive communication according to changing traffic profile is possible by the use of time division duplexing (TDD), dynamic spectrum sharing and flexible slot structures. Moreover, 5G uses SDN and NFV within the core network in case there is a need to provide programmable and scalable services which are very crucial in industrial applications where there is need to deploy and orchestrate services quickly.

4.2 Edge AI Capabilities

When combined with smart data processing technologies, edge computing has great benefits in industrial settings that must be quickly low-latency and very reliable. Edge-based analytics involves computation, decision and control tasks operating close to data source, usually embedded systems, gateways or industrial controllers and thus not requiring the transmission of data to centralized cloud locations. Some of the industrial considerations that this architectural shift deals with are the need to have less latency, better privacy, better use of bandwidth and a more autonomous system [22-23]. Ability to run the inference and the pattern recognition algorithm at the edge devices is one of the fundamental abilities of edge intelligence. Such algorithms are normally trained not in real time, so large datasets are required, and such algorithms are gradually deployed in an optimized version. Edge intelligence can be utilized in industries, where the information on sensors attached to any equipment (motors, pumps, turbines) can be analyzed to streamline predictive maintenance of anything. Edge devices can predict probable issues in advance before they may happen by recognizing peculiarities in patterns in vibration, temperature or acoustic signals, reduce unexpected downtimes and enhance asset utilization.

Real-time control optimization is another useful ability. Edge nodes may easily adjust control actions, without much delay, in order to monitor parameters of operation in processes that require robotic arms, CNC machines, or even conveyor systems. That is essential to applications in which closed-loop control is needed in response times that can be in the sub-millisecond. Also, edge devices can be used to apply the industrial vision system, so products on assembly lines can be checked quickly, and the visual information is processed, and defects can be categorized independently of cloud reliance [24]. Adaptive workload distribution is as well boosted by edge intelligence. It allows priority to be set on system states, network conditions or on levels of criticality of applications. To use an example, in a smart manufacturing setting, only major anomalies or summarized knowledge about the environment must be sent higher-level systems whereas the more minor data could be computed locally with a lightweight computational burden. Besides, the edge nodes provide improved security provisions since they limit the exposure of data outside the regional system. Sensitive industrial data is kept in the working scope so the danger of data breach gets minimized. Access control capabilities, local storage encryption, and implementation of secure boot can be done in edge computing platform to enhance industrial security.

Edge intelligence makes it possible to perform data-driven decentralized, responsive, and scalable operation that best suits the dynamic needs of the smart industries. With the connectivity capabilities of 5G, the same capabilities enable real-time decision-making, resource management, and improved processes reliability in distributed industrial infrastructures.

5. System Architecture

Scalable and well-defined system architecture is needed to deploy real-time industrial intelligence that can be achieved by the utilization of 5G and edge-based data processing. This architecture is usually composed of many layers, each layer is responsible to perform specific computations and data flow as well as functions. These architectural layers are the Device Layer, Edge Layer, Network Layer, and Cloud Layer that, altogether, provide a high degree of non-disruption to the sensing, communication, computation, and control when they are used in conjunction with smart industrial systems. Figure 6. shows the edge computing architecture.

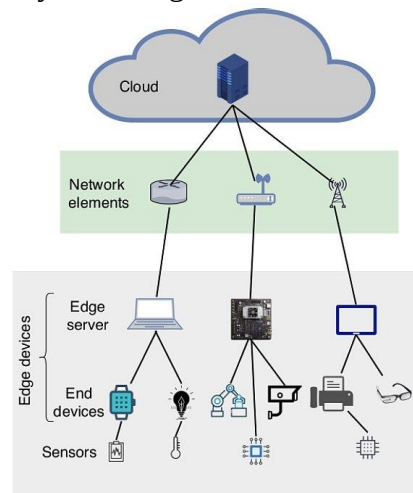


Figure 6: Edge Computing Architecture

Device Layer: The lowest layer is known as the Device Layer and comprises the physical equipment, e.g. sensors, actuators, programmable logic controllers (PLCs), embedded controllers and industrial robots. These are the devices that can gather raw performance data such as temperature, pressure, motion, vibration and machine states. The controls of this layer send signals to actuators so as to make physical changes in accordance to instructions processed. Devices in this layer tend to be fitted with communication modules capable of 5G NR or industrial Ethernet communications which are used to exchange data with the higher layers with high speed and reliability. Of interest in the design on this level are power efficiency, environmental ruggedness and real time I/O interfaces.

Edge Layer : Edge Layer is located near the devices layer usually in a same building or in-house network. It is made up of edge servers, gateways, industrial PC, embedded computing platforms that could perform on local processing functions. It does basic data filtering, signal processing, real-time analytics and executes control logic. It has the ability to run localized decision-making algorithms that respond to critical instances of time, because of the low latencies. The Edge Layer avoids auto loading of the network as it pre-processes and aggregates the data upon sending each to centralized resources. It supports high availability of control systems with partial or unreliable cloud connectivity as well.

Network Layer : It can include the communication infrastructure across and between the device, edge, and cloud parts. This layer can be constructed in a smart industrial environment over 5G networks which incorporate functionalities like URLLC of control signals of critical nature, eMBB of high-throughput applications, and the mMTC to connect large numbers of devices. QoS management, dynamic bandwidth allocation and network slicing support the network layer in guaranteeing that each industrial application is provided with the required

degree of service performance. It also incorporates transport protocols, as well as security measures that protect transport time-sensitive and confidential information.

Cloud Layer : The Cloud Layer is a center where long-term data are stored on a large-scale computation, historical trends analysis, and enterprise-level system integration. It offers largescale model training, distributed assets management, and supervisory control efforts. The cloud tier integrates with enterprise resource planning (ERP), manufacturing execution systems (MES) and supply-chain platforms to support end-to-end visibility within the operations. It also facilitates introducing updates as well as global coordination of various industrial locations with the help of centralized dashboards.

The strength here is that this is a multi layered system architecture that gives good distribution of the resources and also is scalable. Mobile computation involves assigning tasks needing to be responded to at the local scale to the devices or edges, and other tasks that involve high amounts of computation or use or long-term analysis to the cloud infrastructures. Layer aggregation based on low-latency and high-reliability connections in 5G is the crux of time-sensitive industry intelligence in the next-generation smart factory.

6. Industrial Applications

The integration of 5G communication systems with edge-based data processing technologies is revolutionizing industrial environments by enabling real-time, reliable, and decentralized decision-making. The high-speed connectivity, ultra-low latency, and massive device support provided by 5G networks, together with the localized computational capabilities of edge computing, facilitate efficient handling of time-critical and bandwidth-intensive applications. This synergy enhances operational efficiency, productivity, safety, and intelligent automation in smart industries. Moreover, processing data closer to the source minimizes communication delays, reduces bandwidth usage, and decreases dependence on centralized cloud infrastructures. Consequently, the combination of 5G and edge computing has emerged as a fundamental technology for supporting advanced applications such as industrial IoT, smart manufacturing, autonomous systems, healthcare monitoring, and intelligent transportation systems [25]. 5G technology provides significant advantages such as high data rates, low latency, massive device connectivity, and improved reliability; however, its performance may not be uniform across all industrial applications. The effectiveness of 5G depends on several factors, including network architecture, spectrum allocation, infrastructure availability, environmental conditions, application requirements, and device capabilities. For example, applications such as autonomous robotic control, real-time monitoring, and industrial automation may benefit greatly from Ultra-Reliable Low-Latency Communication (URLLC), whereas applications involving large-scale sensor deployments may primarily utilize massive Machine-Type Communication (mMTC) features. In practical industrial environments, factors such as electromagnetic interference, physical obstructions, network congestion, and limited coverage in remote areas can affect the actual performance of 5G systems. In addition, some industrial operations may still rely on hybrid communication architectures that combine 5G with Wi-Fi, Ethernet, TSN, or legacy industrial networks to meet specific operational requirements. Therefore, it is important to avoid generalizing 5G capabilities across all industrial scenarios and instead evaluate its suitability based on the specific needs and constraints of each application.

Predictive Maintenance : One of the most important applications is predictive maintenance as a part of which the real-time equipment health monitoring is provided to foresee premature breakdowns. Edge-enabled vibration sensors, temperature monitors or acoustic detectors will also be used to analyze machinery health on-site and identify problems early

enough before they permanently affect machines. 5G will provide high-bandwidth and low-latency access to only critical alerts or trends are sent to supervising systems, thus minimizing non-essential data and guaranteeing the timely maintenance process can occur.

Automation and Industrial Robotics : Contemporary robots found in the manufacturing industry perform roles like assembly, provision of welding, transportation of materials and packaging. Such operations involve a precise timing and synchronization, and it is facilitated by URLLC capabilities of 5G. Robots in close proximity to edge computing platforms can have real-time control computations and algorithms that perform the path optimization. This configuration is less latent than cloud-based command and therefore ideal in operations of the robot, collision avoiding and responsive performance of the robots in complex production lines.

Real Time Quality Inspection : Industrial cameras and edge image processing units perform visual inspection in high speed, production line-based systems, to locate defects on the surfaces, dimensional defects, or assembly defects. When using eMBB and Edge Processing, high-resolution information data can be processed in real-time without the need to use cloud resources. This will enable instantaneous feedback to machinery to help it take corrective measures and improve the quality of the products and the scrap percentage.

Self-Driving Autonomous Vehicles (AGVs) : Mobile robots and AGV are well known throughout warehouses and shop floors to carrying material. Such cars depend both on constant communication with control systems and sensor fusion onboard, allowing them to navigate. 5G is expected to provide steady connectivity, and edge computing is likely to help in real-time navigation, onboard route planning, and obstacle detection. This configuration allows AGVs to handle dynamic environments whose layouts change or have high-traffic situations [26].

Control Process and Remote Monitoring : Centralized monitoring and control of distributed assets is important in oil and gas industries, power generation industries and chemical processing industries. Using 5G backhaul, edge IoT devices in remote stations have time to analyze local sensor data and use predefined control logic to make decisions at remote locations. Remote supervision, alarm processing, and regulatory compliance are possible with a low-latency connection.

These kinds of applications can be made because the integration of 5G and edge computing means a more reliable and higher throughput, and agile operations at the industrial domains. This convergence serves a critical role in the attainment of objectives of Industry 4.0 by facilitating the flexible production systems, smart supply chain, and robust infrastructures.

7. Limitations and Challenges

Even with great potential that 5G and edge-based data processing have in changing the nature of industrial processes, there are a number of technical and operational barriers in its adoption that limit its widespread use in real-time industrial settings. These constraints are in infrastructure, interoperability, security, scalability as well as system integration. Figure 7 shows the challenges and potential applications of edge intelligence.

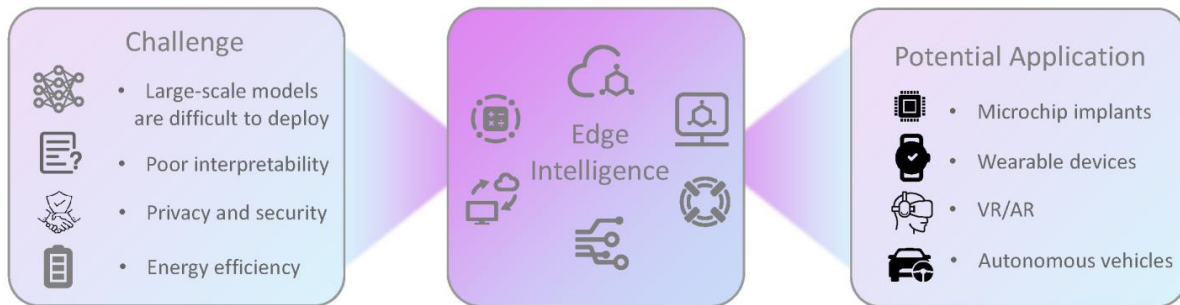


Figure 7: Challenges and potential applications of edge intelligence [18]

Despite the significant advantages of integrating 5G and Edge AI in industrial systems, several limitations and challenges still exist. One major challenge is maintaining ultra-low latency and reliable communication in large-scale industrial environments where network congestion, signal interference, and infrastructure limitations may affect performance. Security and privacy concerns also remain critical because distributed edge nodes and connected devices increase the attack surface for cyber threats, unauthorized access, and data breaches. In addition, deploying and managing Edge AI systems require high computational resources, advanced hardware, and continuous maintenance, which increase implementation costs and energy consumption. Another limitation is interoperability among heterogeneous devices, communication protocols, and vendor-specific platforms, making seamless integration difficult in complex industrial ecosystems. Furthermore, Edge AI models may face constraints related to limited processing power, storage capacity, and real-time model updates at edge nodes. Scalability and fault management also become challenging as the number of connected devices increases. Therefore, further research is needed to develop standardized architectures, energy-efficient edge platforms, secure communication mechanisms, and intelligent resource management techniques to improve the reliability and practical deployment of 5G-enabled Edge AI systems in industrial applications.

Network Infrastructure Implementation : Existing the 5G networks and deployment especially in indoor industrial environment need dense small cell deployment especially in indoor industrial sites, high-frequency cell stations as well as strong backhaul backbones. The adoption of the millimeter-wave bands, which are useful at very high data rates, has the respective disadvantage of low penetrations as well as vulnerability to physical barriers like walls and machineries. Creating coverage in the world of industrial premises that is competent and can be recognized on a regular basis is a considerable difficulty and challenge that will demand planning and a vast sums of money.

Integration Complexity : The problem with edge computing systems is that they can be mixed with the existing industrial automation equipment and control systems that are sometimes legacy protocols, proprietary devices, and heterogeneous platforms. Compatibility between the latest edge gateways and the existing programmable logic controllers (PLCs), supervisory control systems, and sensors requires customized adaptation layers and middleware solutions, which add complexity to deployment.

Latency and synchronization : Although 5G is having ultra-low latency on paper, deterministic timing in industrial control applications is a problem in reality. Timing variation may be added by factors like network congestion, interference with signals and switching delay. In applications where synchronization on a sub-millisecond time frame is important (e.g. multi-axis robot coordination or high-speed motion control), this can be a problem.

Security And Data Privacy : Although edge deployments minimize the transfer of data to distant servers, new attack vectors are associated with edge deployments. Without proper authorization, it is possible to gain unauthorized physical access to edge devices, the vulnerability of firmware, or insecure communication channels and end up with data being lost or system compromised. Establishing secure boot, device authentication and encryption of data transport are necessary yet frequently difficult to perform in large-scale and distributed environments.

Scalability and limit in resource availability : Edge devices are usually operating under various limitations including low computation capacity, memory and energy shortage. The workload distribution and optimization requirements to scale edge intelligence on a large number of devices is efficient. Besides that, it is difficult to keep uniform performance with geographically distributed edge nodes in the aspects of configuration, monitoring, and updates.

The integration of 5G and edge computing provides significant technical benefits for real-time industrial intelligence. However, it is essential to address the associated technical challenges to ensure reliability, security, and scalability in industrial environments.

8. Future Directions

The integration of 5G communication and edge-based data processing in industrial environments is still evolving, and future developments are expected to enhance the performance, reliability, and scalability of these technologies. A major advancement is the integration of Time-Sensitive Networking (TSN) with 5G, where the Ultra-Reliable Low-Latency Communication (URLLC) capabilities of 5G can support deterministic communication, bounded latency, and precise synchronization required for mission-critical applications such as robotic coordination and distributed motion control. Another important research direction is the development of energy-efficient edge computing platforms capable of operating in harsh and remote industrial environments with optimized power consumption and real-time processing. The adoption of private 5G networks is also expected to grow in industries requiring customized Quality of Service (QoS), enhanced security, and greater operational control. Future industrial systems may rely on hybrid architectures integrating public and private 5G networks managed through unified orchestration platforms. In addition, advancements in secure over-the-air (OTA) updates, device onboarding, and distributed fault management will support large-scale deployments. The integration of Low Earth Orbit (LEO) satellite communication with edge systems may further extend connectivity to remote industrial sectors such as mining, maritime logistics, and energy distribution, enabling more autonomous, adaptive, and resilient industrial operations.

9. Conclusions

The choice of combining 5G communication systems with edge-based data handling is transforming industrial processes by providing real-time knowledge, small-scale control, and automatic scaling. Such convergence has been aided in the major industries including predictive maintenance, robotic automation, quality inspection and distant asset monitoring. With the advantage of eMBB, URLLC, and mMTC functions and by localized edge processing, increased productivity and operational safety with responsiveness can be attained in the industries. Nevertheless, the problem of infrastructure, interoperability, precision of latency, and security are still the dimensions that need future investigation and standardization. Further development

of edge equipment architecture, orchestration of networks, and system integration platforms will also be critical in providing completely interconnected and smart industrial systems.

Acknowledgment: Not Applicable.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study

References

- [1] K. K. Vaigandla and N. Venu, "A survey on future generation wireless communications-5G: Multiple access techniques, physical layer security, beamforming approach," *J. Inf. Comput. Sci.*, vol. 11, no. 9, pp. 449–474, 2021.
- [2] Z. Li, "5G URLLC: Design challenges and system concepts," in *Proc. 2018 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2018.
- [3] K. K. Vaigandla, "Communication technologies and challenges on 6G networks for the Internet: Internet of Things (IoT) based analysis," in *Proc. 2022 2nd Int. Conf. Innovative Practices Technol. Manage. (ICIPTM)*, vol. 2, 2022.
- [4] B. Liang, M. A. Gregory and S. Li, "Multi-access edge computing fundamentals, services, enablers and challenges: A complete survey," *J. Netw. Comput. Appl.*, vol. 199, 2022.
- [5] T. Nguyen, H. Nguyen and T. N. Gia, "Exploring the integration of edge computing and blockchain IoT: Principles, architectures, security, and applications," *J. Netw. Comput. Appl.*, vol. 226, 2024.
- [6] Y. O. Imam-Fulani, "5G frequency standardization, technologies, channel models, and network deployment: Advances, challenges, and future directions," *Sustainability*, vol. 15, no. 6, 2023.
- [7] S. Senger and P. K. Malik, "A comprehensive survey of massive-MIMO based on 5G antennas," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 32, no. 12, 2022.
- [8] K. K. Vaigandla and N. Venu, "Survey on massive MIMO: Technology, challenges, opportunities and benefits," 2021.
- [9] R. Dangi et al., "Study and investigation on 5G technology: A systematic review," *Sensors*, vol. 22, no. 1, Art. no. 26, 2022.
- [10] S. Mattisson, "An overview of 5G requirements and future wireless networks: Accommodating scaling technology," *IEEE Solid-State Circuits Mag.*, vol. 10, no. 3, pp. 54–60, 2018.
- [11] A. Tikhomirov, E. Omelyanchuk, and A. Semenova, "Recommended 5G frequency bands evaluation," in *Proc. 2018 Syst. Signals Generating Processing Field Board Commun.*, 2018.
- [12] K. K. Vaigandla, "A systematic survey on artificial intelligence in 6G wireless networks: Security, opportunities, applications, advantages, future research directions and challenges," *Babylonian J. Artif. Intell.*, pp. 99–106, 2025.
- [13] J. Lu et al., "5G enhanced service-based core design," in *Proc. 2019 28th Wireless Opt. Commun. Conf. (WOCC)*, 2019.
- [14] M. Zieba, M. Natkaniec and P. Borylo, "Cloud-enabled deployment of 5G core network with analytics features," *Appl. Sci.*, vol. 14, no. 16, 2024.
- [15] M. Attaran, "The impact of 5G on the evolution of intelligent automation and industry

-
- digitization,” *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 5, pp. 5977–5993, 2023.
- [16] P. Orosz, “QoS guarantees for Industrial IoT applications over LTE-A: A feasibility study,” in *Proc. 2019 IEEE Int. Conf. Ind. Cyber Phys. Syst. (ICPS)*, 2019.
- [17] Y. Xu, “Artificial intelligence: A powerful paradigm for scientific research,” *The Innovation*, vol. 2, no. 4, 2021.
- [18] T. Wang, “Deploying AI on edge: Advancement and challenges in edge intelligence,” *Mathematics*, vol. 13, no. 11, Art. no. 1878, 2025.
- [19] R. Singh and S. S. Gill, “Edge AI: A survey,” *Internet Things Cyber-Phys. Syst.*, vol. 3, pp. 71–92, 2023.
- [20] K. K. Vaigandla, S. Bolla and R. Karne, “A survey on future generation wireless communications-6G: Requirements, technologies, challenges and applications,” *Int. J.*, vol. 10, no. 5, pp. 52021, 2021.
- [21] J. Cheng, “5G in manufacturing: A literature review and future research,” *Int. J. Adv. Manuf. Technol.*, vol. 131, no. 11, pp. 5637–5659, 2024.
- [22] S. S. Gill, “Edge AI: A taxonomy, systematic review and future directions,” *Cluster Comput.*, vol. 28, no. 1, 2025.
- [23] J. Shao and X. Li, “AI flow at the network edge,” *IEEE Netw.*, 2025.
- [24] M. Elahi, “A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment,” *Discover Artif. Intell.*, vol. 3, no. 1, 2023.
- [25] M. Ergen et al., “Edge computing in future wireless networks: A comprehensive evaluation and vision for 6G and beyond,” *ICT Express*, vol. 10, no. 5, pp. 1151–1173, 2024.
- [26] M. Aizat, A. Azmin and W. Rahiman, “A survey on navigation approaches for automated guided vehicle robots in dynamic surrounding,” *IEEE Access*, vol. 11, pp. 33934–33955, 2023.