Power-Aware Video Transmission in 5G Telemedicine: Challenges, Solutions, and Future Directions

Qiuhong SHI, Mingjing CAO*, Yun BAI
College of Physics and Information Engineering, Cangzhou Normal University, Cangzhou 061000, China

***Abstract*—The transformation of healthcare delivery through telemedicine has been significantly accelerated by the deployment of 5G networks and the integration of Internet of Things (IoT) technologies. These advancements enable real-time video-based medical services, including remote operations, urgent care, mobile procedures, and virtual consultations. However, the energy limitations of IoT devices present significant questions regarding long-term video quality and system efficiency. This survey reviews state-of-the-art solutions envisioned for enhancing video transmission in telemedicine environments that are limited by energy consumption but made possible by 5G connectivity. The study discusses recent advances, including efficient video compression techniques, computation offloading via edge computing, adaptive streaming procedures, and dynamic 5G architecture-aware resource scheduling, such as network slicing. It discusses the trade-off between power efficiency and video performance for various telehealth scenarios. Using a scenario-based analysis with a unifying integration framework, this work advances research into energy-efficient e-health systems.**

***Keywords*—Telemedicine; Internet of Things; energy efficiency; video transmission**

I. Introduction

The World Health Organization (WHO) recognizes telemedicine as a form of medical treatment delivered through audiovisual and digital communications [1]. Medical care services include diagnosis, consultation, treatment, health education, and the transmission of medical records. The telemedicine architecture can be divided into three layers: sensing, gateways, and healthcare facilities. The Internet of Things (IoT) is a network of people connected to everything, anytime, anywhere. IoT is defined differently in the literature. The IoT provides various healthcare services, including fitness programs, remote health tracking, rehabilitation, and elder care [2]. Telemedicine is essential for remotely monitoring treatments and medications at home [3]. The IoT architecture for telemedicine is based on a range of medical devices, sensors, and imaging devices. Therefore, IoT and telemedicine will contribute to the future comprehensive remote monitoring of the healthcare industry.

Telemedicine has revolutionized healthcare by allowing long-distance consultations, monitoring, and diagnosis through video-based interactions [4]. Along with telemedicine applications that continue to expand exponentially, trustworthy and energy-aware video transmission has become essential, particularly in applications that include battery-constrained IoT technologies, such as wearable health sensors and remote diagnostic tools [5]. These tend to have limited battery life and restricted processing capabilities, for which energy efficiency must be a key consideration. The emergence of 5G networks offers potential solutions to these issues through low latency, higher bandwidth, and highly reliable connectivity [6].

While there is an increasing amount of literature addressing energy efficiency in 5G and IoT systems, previous reviews tend to overgeneralize the solution without adapting it to the specific requirements of telemedicine applications. Most research fails to consider the particular issues related to continuous video transmission in cases such as telemedicine applications in remote surgery, emergency procedures, and long-term patient monitoring via battery-constrained IoT devices. Our review addresses this gap by providing a real-case-driven, holistic treatment of energy-efficient video transmission schemes specifically crafted for telemedicine. By integrating technical solutions with actual healthcare applications, this paper offers focused insights that are often lacking in more general, non-specialized surveys.

This article examines recent advances in energy-efficient video transmission technologies for telemedicine, focusing on the role of 5G and IoT devices. We examine the technical challenges, assess existing solutions, and outline future research directions to enhance energy efficiency in this area. This article makes the following key contributions:

- An in-depth investigation of cutting-edge video encoding and adaptive streaming strategies is provided, aiming to reduce energy consumption in telemedicine applications that utilize resource-constrained IoT devices.
- The role of 5G technology in enhancing video transmission quality and improving energy efficiency in telemedicine applications is demonstrated through network slicing, edge computing, and dynamic power control.
- An analysis of conflicts between Quality of Service (QoS) and energy efficiency is presented, addressing the interoperability issue and identifying the need for AI-driven solutions and energy-conscious protocols in the future development of telemedicine, coupled with IoT and 5G.

II. Background and Motivation

This section outlines the requirements for telemedicine and video transmission, the contribution of 5G networks to the advancement of telemedicine, and the challenges posed by resource-limited IoT devices in healthcare.

A. Telemedicine and Video Transmission Requirements

Telemedicine applications rely on real-time video transmission for remote consultations, diagnosis, and patient monitoring [7]. Video transmission quality and efficiency are crucial for delivering seamless service to both patients and healthcare providers. This subsection outlines the minimum requirements for telemedicine video transmission, including video quality, latency, bandwidth, and energy efficiency, with consideration for resource-constrained IoT devices. Video quality is crucial for telemedicine, as doctors rely heavily on visual information to diagnose and monitor patients [8]. The quality of a video is usually measured by its bit rate, frame rate, and resolution, as shown in Table I.

Telemedicine requires real-time interaction between healthcare professionals and patients. High latency can negatively impact communication and diagnostic processes, as illustrated in Fig. 1, which shows the increasing latency associated with higher video resolutions [9]. In telemedicine, real-time applications require a latency of under 100 milliseconds. Providing high-quality video without interruptions during telemedicine transmissions requires sufficient bandwidth. As shown in Table II, the network type (e.g., Wi-Fi, 4G, 5G) and the number of concurrent users affect the available bandwidth. Typically, bandwidth requirements for telemedicine applications involving high-definition video range from 1 Mbps to 5 Mbps.

Energy efficiency is crucial, particularly when utilizing limited-life battery-powered IoT devices. Telemedicine is one of the most energy-intensive operations in video transmission, involving uninterrupted data encoding, transmission, and reception. Minimizing energy consumption is crucial for enhancing device longevity and maintaining device functionality [10]. Table III depicts energy consumption for different telemedicine video encoding formats. Energy consumption versus telemedicine video quality is shown in Fig. 2.

TABLE I. Key Video Quality Parameters for Telemedicine

Parameter	Definition	Typical value in telemedicine
Resolution	Number of pixels in each video frame	720p, 1080p, 4K (depending on use case)
Frame rate	Number of frames per second (fps)	30 fps (for real-time communication)
Bit rate	Amount of data transmitted per second	1-3 Mbps (depending on network conditions)

TABLE II. Bandwidth Requirements for Telemedicine Video Transmission across Different Network Types

Network type	Bandwidth availability	Suitability for telemedicine
Wi-Fi (802.11n)	50-300 Mbps	High-definition video
4G LTE	5-12 Mbps	Basic telemedicine applications
5G (eMBB)	100-1000 Mbps	Ultra-high-definition video

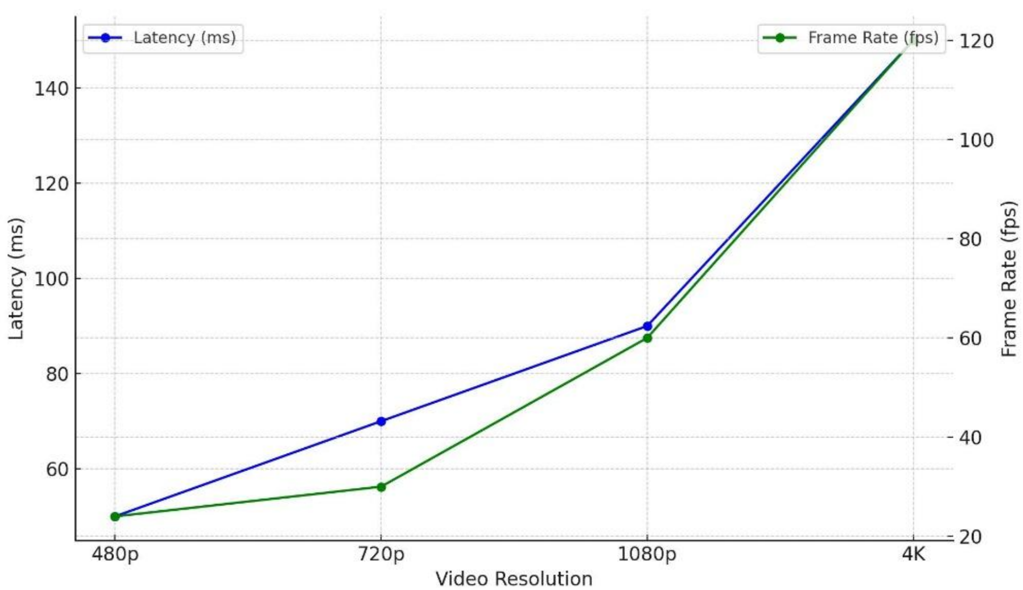

Fig. 1. Relationship between video quality and latency in telemedicine applications.

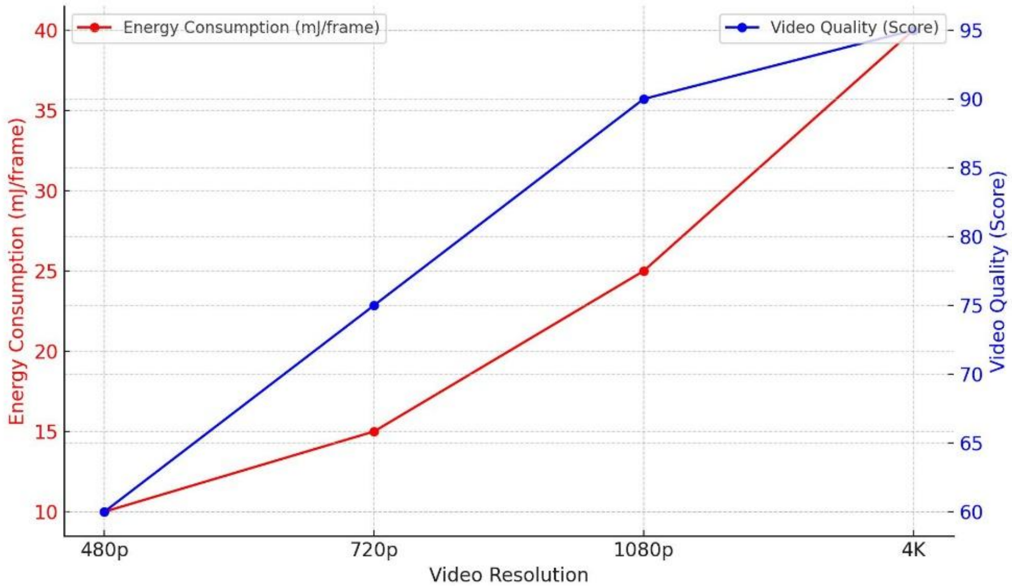

Fig. 2. Relationship between video quality and energy usage in telemedicine.

TABLE III. ENERGY CONSUMPTION OF DIFFERENT VIDEO ENCODING FORMATS IN TELEMEDICINE

Video encoding format	Energy consumption (mJ/frame)	Efficiency
H.264/AVC	15-30	Moderate
H.265/HEVC	8-20	High
VP9	10-25	Moderate

B. 5G Networks and Telemedicine

The 5G network is a transformative technology for telemedicine, offering features such as ultra-low latency, high-speed data transmission, and enhanced interoperability. These features facilitate real-time video transmission in telemedicine applications, facilitating secure and effective interaction between patients and healthcare professionals [11].

5G networks also possess various technologies that qualify them for telemedicine services, including Massive Machine-Type Communications (mMTC), Ultra-Reliable Low Latency Communication (URLLC), and Enhanced Mobile Broadband (eMBB) [12]. eMBB enables bandwidth-intensive applications such as streaming ultra-high-definition video. URLLC provides minimal delay in communication, enabling real-time discussion between patients and medical specialists. mMTC provides opportunities for numerous IoT devices in telemedicine scenarios that require multiple attached sensors, wearables, and remote diagnostics. As highlighted in Table IV, 5G offers specific benefits to telemedicine over previous generations of mobile networks.

High-resolution video streaming makes tele-diagnostics possible with great accuracy. High-bandwidth and low-latency 5G technology would allow real-time teleconsults without delay [13]. IoT-capable health devices, such as wearables and sensors, can collect real-time data on a patient's health status. The mMTC feature of the 5G network can enable the mass deployment of such devices, allowing healthcare professionals to receive real-time data with which to monitor patients. Augmented Reality (AR) technologies are integral to 5G's URLLC technology, especially in remote surgery, where precision and real-time feedback are crucial for telemedicine. The 5G network introduces sleep modes and energy-aware protocols to increase device longevity. Table V illustrates how 5G's energy-efficient protocols impact IoT devices in telemedicine.

TABLE IV. COMPARISON OF 5G WITH PREVIOUS MOBILE NETWORK GENERATIONS FOR TELEMEDICINE

Feature	3G	4G	5G	Impact on telemedicine
Latency	~100-200 ms	~30-50 ms	~1-10 ms	Enables real-time and low-latency video consultations
Bandwidth	~2 Mbps	~10-50 Mbps	~100 Mbps-1 Gbps	Supports high-definition video streaming for remote diagnosis
Device connectivity	Limited	Moderate	Massive (up to 1 million/km²)	Connects large numbers of IoT devices for continuous monitoring
Energy efficiency	Low	Moderate	High	Extends the battery life of IoT devices used in telemedicine
Reliability	Moderate	High	Ultra-reliable	Critical for life-saving remote interventions and consultations

TABLE V. IMPACT OF 5G ENERGY-EFFICIENT PROTOCOLS ON IOT DEVICES IN TELEMEDICINE

Technology	Energy efficiency mechanism	Impact on IoT devices
5G network slicing	Isolating critical applications into slices and optimizing resource use	Reduced energy consumption for telemedicine services
Energy-aware scheduling	Adjusting power consumption based on device activity	Extended battery life of wearable devices in telemedicine
Sleep modes	Permitting devices to sleep when not transmitting	Enhanced energy efficiency, especially in continuous monitoring devices

C. Resource-Constrained IoT Devices in Telemedicine

Telemedicine relies heavily on low-resource devices, such as wearables, portable diagnostic instruments, and sensors. These devices are typically limited in terms of processing power, memory, and battery life, posing significant challenges to energy efficiency and reliability [14]. Innovative approaches to power management, communication standards, and data transmission strategies are needed to ensure the efficient integration of these devices into telemedicine systems.

The lightweight and portable characteristics of IoT devices designed for telemedicine impose some limitations. Most IoT components operate on batteries, limiting their ability to monitor or transmit for long periods. These devices typically have low computing resources, making advanced data processing difficult. Data loss is a significant problem when storage capacity is limited, so using efficient compression and transmission strategies is essential. Table VI summarizes the general limitations of IoT devices used in telemedicine [15].

Resource-constrained IoT devices in telemedicine face several challenges. Video transmission, real-time monitoring, and data collection from IoT devices consume a significant amount of power [16]. Telemedicine environments often require frequent recharging or early device failure, affecting the continuity of care. Delays in IoT devices operating over low-power wireless communication standards, such as Zigbee and Bluetooth, can affect telemedicine services. Table VII compares energy consumption for standard IoT communication protocols.

TABLE VI. RESOURCE-CONSTRAINED IOT LIMITATIONS USED IN TELEMEDICINE

IoT device type	Battery life	Processing power	Memory size	Communication range
Wearable health monitor	1-3 days (continuous use)	Low (single-core MCU)	256 KB – 1 MB	10-30 meters (Bluetooth)
Portable diagnostic tool	6-12 hours (per charge)	Moderate (dual-core CPU)	4-8 GB	50-100 meters (Wi-Fi)
Implantable sensor	Several months (low use)	Very Low (minimal CPU)	64-128 KB	5-10 meters (NFC)

TABLE VII. AN OVERVIEW OF COMMON IOT COMMUNICATION PROTOCOLS IN TELEMEDICINE

Protocol	Typical range	Data rate	Power consumption	Use case in telemedicine
Bluetooth Low Energy (BLE)	10-30 meters	1 Mbps	Low	Wearable health monitors, fitness trackers
Zigbee	10-100 meters	250 Kbps	Very Low	Home health monitoring systems
Wi-Fi	50-100 meters	100-1000 Mbps	High	Video consultations, portable diagnostic tools
NFC	4-10 centimeters	< 500 Kbps	Minimal	Implantable medical sensors

III. ENERGY-EFFICIENT VIDEO TRANSMISSION STRATEGIES

This section discusses energy-efficient video transmission strategies, including video encoding techniques to reduce data sizes, adaptive bitrate streaming for dynamic quality adjustments, and edge computing with caching.

A. Video Encoding

In telemedicine applications, video encoding is crucial for reducing the size of video streams while maintaining quality. Resource-constrained IoT devices and 5G-enabled telemedicine systems require efficient video encoding, which impacts bandwidth usage, transmission latency, and energy consumption [17]. Video encoding reduces the amount of data required for transmission by compressing video files. This process aims to optimize network resources and minimize the load on IoT devices. Standard techniques include H.264/AVC (advanced video encoding), H.265/HEVC (high-efficiency video encoding), and VP9.

H.264 is a popular video compression standard because it has a high compression ratio and is compatible with a wide range of devices [18]. H.265 has better compression than H.264, reducing the bitrate required to transmit high-resolution video by approximately 50% [19]. Google developed VP9 as a royalty-free, open-source video codec that YouTube popularly uses for streaming videos. VP9 is similarly efficient as H.265 but is royalty-free [20]. Table VIII compares the performance and compression efficiency of these encoding standards.

TABLE VIII. PERFORMANCE AND COMPRESSION EFFICIENCY COMPARISON OF COMMON VIDEO ENCODING STANDARDS

Encoding standard	Compression ratio	Energy consumption (mJ/frame)	Video quality retention	Typical use in telemedicine
H.264/AVC	Moderate	15-30	High	Real-time video consultations
H.265/HEVC	High	8-20	Very high	High-definition video diagnosis
VP9	High	10-25	High	Low-latency video transmission

TABLE IX. ENERGY CONSUMPTION OF VIDEO ENCODING STANDARDS AT DIFFERENT RESOLUTIONS USED IN TELEMEDICINE

Resolution	H.264/AVC (mJ/frame)	H.265/HEVC (mJ/frame)	VP9 (mJ/frame)	VP9 (mJ/bit)
480p	10-15	5-8	7-10	~0.04
720p	15-20	8-12	10-15	~0.035
1080p	20-25	12-16	15-20	~0.03
4K	25-30	18-20	28-32	~0.025

Energy efficiency in video encoding is crucial for telemedicine applications, particularly in resource-constrained IoT devices [21]. Efficient video coding methods reduce the computing load on devices, minimize battery consumption, and reduce energy consumption for video transmission over networks. Compared to H.264, H.265 offers the same video quality at half the bit rate, significantly reducing the amount of data transferred. This leads to lower energy consumption. Although VP9 provides similar compression for H.265, it is particularly beneficial for applications that require open-source codecs. Although it uses slightly more power than H.265, it remains a viable option for energy-efficient video transmission. As shown in Table IX, different video coding standards in telemedicine consume different amounts of energy according to their resolution.

Video quality and energy consumption must be optimally balanced in telemedicine. In adaptive video encoding, bit rate, resolution, and frame rate are dynamically adjusted based on real-time network conditions and the energy availability of the transmitting device. This approach is ideally suited for telemedicine consultations and remote monitoring.

B. Adaptive Bitrate Streaming

Adaptive Bitrate Streaming (ABR) provides dynamic adjustments to video quality in response to real-time network conditions and device capabilities in telemedicine [22]. The technology ensures video streams are delivered with minimal buffering and optimal quality, even in environments with fluctuating bandwidth. Video quality and energy efficiency are essential in telemedicine, especially with resource-constrained IoT devices and wireless networks.

ABR encodes video content at multiple bitrates and resolutions. Client devices (e.g., smartphones or IoT devices) constantly monitor network conditions during video playback to ensure smooth streaming. By dynamically adjusting quality levels based on bandwidth availability, device performance, and battery life, this technology ensures uninterrupted transmission while minimizing energy consumption. Client-side monitoring, multiple encoded streams, and dynamic customization are crucial elements of ABR. Videos are encoded at different bit rates to accommodate various network speeds and device capacities (e.g., 360p, 720p, 1080p). The client device sets the video bitrate based on network bandwidth and battery life. Video quality is automatically adjusted according to real-time conditions to ensure smooth playback.

ABR offers several advantages in telemedicine applications, including reduced buffering, improved energy efficiency, and an optimized user experience [23]. A smooth and continuous video stream is ensured by dynamically adjusting the video quality to match the available bandwidth. IoT devices with resource constraints can conserve energy by reducing their bitrate and resolution when network conditions are poor, thereby preventing them from processing high-quality streams that can drain their batteries. Patients and healthcare providers can experience uninterrupted video consultations, even in low-bandwidth or fluctuating network conditions. Table X compares the different video bitrates applicable to telemedicine with their bandwidth requirements and energy consumption.

TABLE X. VIDEO BITRATES, BANDWIDTH, AND ENERGY CONSUMPTION IN TELEMEDICINE USING ABR

Resolution	Bitrate (Kbps)	Bandwidth requirement	Energy consumption (mJ/frame)	Use in telemedicine
360p	500-1000	Low	5-10	Basic consultations
720p	1000-3000	Medium	10-20	Standard remote diagnostics
1080p	3000-6000	High	20-25	High-quality video consultations
4K	6000+	Very High	25-30	High-definition telemedicine

C. Edge Computing and Caching

Telemedicine applications benefit significantly from edge computing and caching, two critical technologies in resource-constrained environments, such as rural areas and mobile health services. These emerging technologies present computing resources proximal to IoT devices and users, consequently entailing low latency, reduced bandwidth usage, and lower energy utilization [24].

With edge computing, data is handled nearer the point of origin (e.g., IoT devices) than from distant cloud servers. This localized processing reduces long-distance data transfers, which is particularly beneficial for telemedicine, where real-time response time is essential. Telemedicine relies heavily on edge computing for real-time video processing, data analysis, AI-driven diagnosis, and latency reduction. Encoding, decoding, and compressing video on the edge reduces processing loads on IoT devices while also conserving battery life. By locally processing medical data at the edge, edge-based AI diagnosis is performed earlier with reduced latency compared to cloud-based systems. Telemedicine services also require minimal latency, as processing data on network edges is crucial for emergency care and real-time consultations. Table XI shows the advantages of edge computing in telemedicine over conventional cloud computing.

TABLE XI. EDGE COMPUTING VS. CLOUD COMPUTING IN TELEMEDICINE

Attribute	Edge Computing	Cloud Computing
Latency	Low (data processed locally)	High (data processed in the distant cloud)
Bandwidth usage	Low (reduced data transmission)	High (frequent data transfer to the cloud)
Energy consumption	Low (offloads computation from devices)	High (IoT devices perform more work)
Real-time responsiveness	High	Moderate
Security	Exposed to edge breaches; AES-256 adds 5–10 mJ/frame	Centralized control; higher encryption scalability

Caching reduces latency and bandwidth usage by temporarily storing frequently accessed content closer to the user. Using edge caching in telemedicine facilitates faster access to medical videos, images, and data by storing them near healthcare providers and patients, thereby reducing the need to repeatedly transmit large files. There are two types of applied caching for telemedicine: Content Delivery Network (CDN) caching and dynamic caching of patient data. CDNs enable healthcare providers to access common healthcare content (e.g., training videos or diagnostic images) more quickly. For smooth interactions, patient data (e.g., diagnostic images or video streams) may be cached on the teleconsultation edge.

While edge computing and caching significantly reduce latency and power consumption, they introduce new security threats. Data processed or stored on edge nodes is more susceptible to unlawful access, physical attacks, or malicious firmware attacks due to its mobility or remoteness. For example, cached patient video streams or images might be vulnerable to data leaks in the event of poor encryption or authentication measures.

Encryption algorithms such as AES-256 provide robust protection for sensitive telemedicine data. However, they incur computational overhead that impacts both latency and energy consumption, particularly in real-time scenarios. Experimental benchmarks indicate that AES-256 encryption adds an average 5–7 ms delay per 1 MB of data and consumes 5–10 mJ/frame in video encoding pipelines, as shown in the updated Table XI. These impacts may be acceptable in high-performance edge nodes but are unsuitable for ultra-low-power IoT devices, such as implantable sensors communicating over NFC.

In such applications, lightweight ciphers like PRESENT, SPECK, or HIGHT become more suitable since they offer good security with minimal processing and energy overhead. PRESENT-80 is an example that has been demonstrated to execute below 0.1 mJ per encryption operation on 8-bit microcontrollers, making it a suitable choice for short-range secure telemetry in implantable telemedicine applications.

D. *TinyML and Split Computing for Video Compression*

The most up-to-date advances in lightweight AI computing and split computation provide unprecedented opportunities for telemedicine energy optimization. Two possible approaches, video compression with TinyML and split computing, extend wearable IoT sensor capabilities while maintaining minimal dependency on robust edge/cloud computing.

On-microcontroller machine learning with limited resources is achieved with TinyML. It is possible to implement lightweight pre-processing and compression of video locally with MCUNet-type models, effectively reducing upstream data sizes considerably. Low-bitrate data compression with minimal info loss or diagnostic features being extracted is achievable with TinyML compression training in wearable health monitor deployment scenarios in healthcare.

For instance, Lin, et al. [25] showed that compression using MCUNet-based compression on ARM Cortex-M processors achieved encoding energy consumption that was 30% less than H.265 while retaining decent accuracy on tasks for clinical classification. This saving is particularly significant in continual monitoring applications, where energy levels directly impact service continuity.

In split computing, the video compression task is partitioned between the IoT device and the edge server. The device performs basic pre-processing (e.g., frame sampling, region-of-interest extraction), while the more intensive encoding steps are offloaded to nearby edge nodes. This division reduces device-side computational load without significantly increasing transmission overhead. Let E_{total} denote total energy usage, split between local computation E_{local} and transmission E_{tx}:

$$E_{total} = E_{local} + E_{tx} \quad (1)$$

Split computing minimizes $E\ l\ o\ c\ a\ l$ E local by moving complex operations off-device and optimizes $E\ t\ x$ E tx via lightweight encoding on-site. Simulation studies show that combining TinyML with split computing can achieve up to 40% total energy savings in wearable diagnostic scenarios, compared to conventional full-device encoding with H.264 or H.265.

These emerging methods complement traditional approaches like ABR and H.265 by offering context-aware compression, real-time adaptability, and lower reliance on high-throughput edge infrastructure, making them ideal for deployment in rural or bandwidth-limited settings.

IV. NETWORK OPTIMIZATION FOR ENERGY EFFICIENCY

This section explores network optimization techniques for energy efficiency in telemedicine, including 5G network slicing, energy-efficient routing protocols, and power control with dynamic resource allocation.

A. *Network Slicing in 5G*

Network slicing, a key component of 5G technology, involves the creation of multiple virtual networks from a shared infrastructure, each customized for specific scenarios and service demands [26]. Telemedicine has numerous applications, from real-time video consultations to telesurgery and patient tracking, and network slicing grants each one its respective bandwidth, latency, and connectivity. The technology also automatically boosts energy efficiency by changing network resources according to telemedicine service demands.

Fig. 3 shows the telemedicine slicing frame. 5G network slicing divides one physical network into multiple virtual slices, each with its own Service-Level Agreement (SLA). Users draw upon resources from that slice that align with their needs. Different slices can be set up for varying measures of performance, such as latency for virtual-real-time video consultations, bandwidth for high-definition video streaming, reliability for vital health data transmission, and energy efficiency for limited devices in the IoT. 5G network slicing optimizes network resources for efficient telemedicine operations without increasing energy consumption. The high-priority slice might be dedicated to remote surgery, which requires high reliability and ultra-low latency. In contrast, the low-priority slice might be allocated to energy-efficient health monitoring devices.

Intelligent health applications can use different network slices, as shown in Fig. 4. This industry requires specialized slices for virtual surgery, smart hospitals, and remote health monitoring. An individual slice can be created for applications requiring high throughput, low latency, and ultra-reliability when performing remote surgery. Remote health monitoring applications and wearable devices can provide secure communication. A dedicated slice can connect medical appliances to a central location.

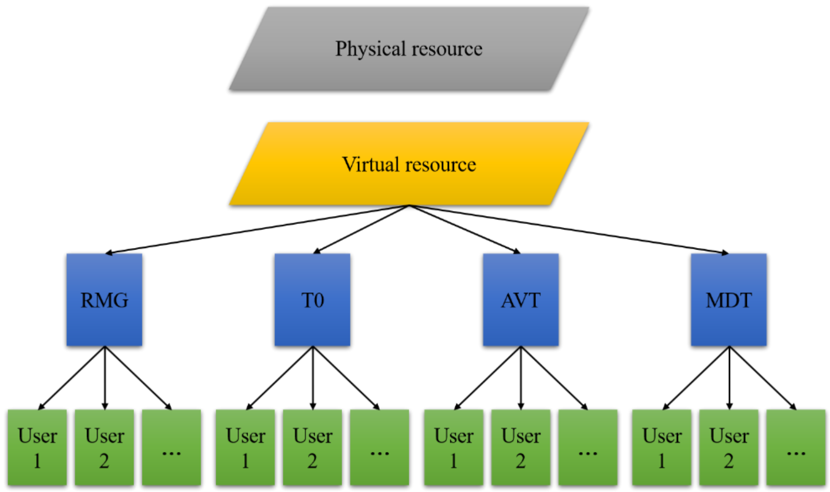

Fig. 3. 5G Network slicing framework for telemedicine.

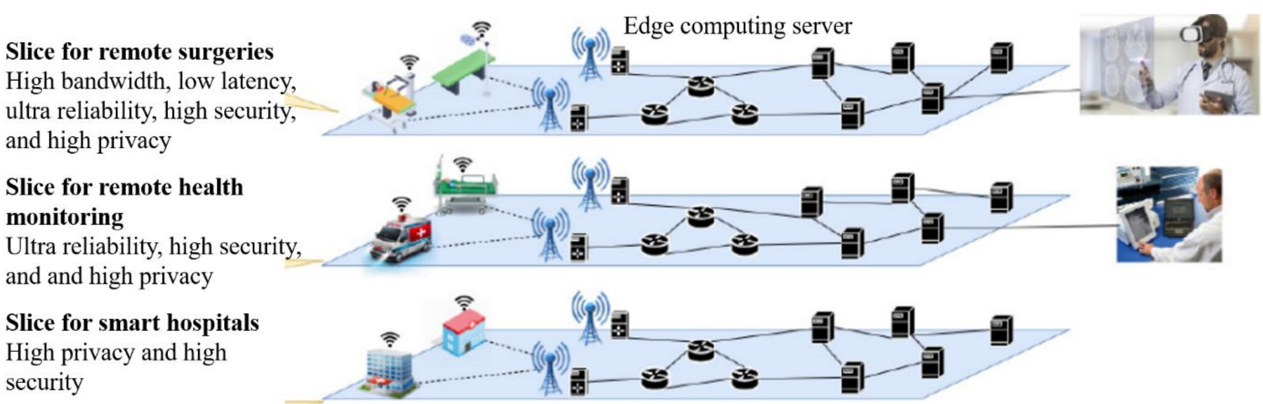

Fig. 4. Network slicing in 5G for telemedicine applications.

As shown in Table XII, network slices can be developed with different energy consumption levels according to application requirements. A high-power slice may be allotted for low-latency video consultations and telecommunication services with high bandwidth for standard real-time connectivity. A low-power slice is ideal for patient telemonitoring, as it reduces bandwidth and latency to achieve low power consumption. Implementing adaptive resource allocation automatically controls energy consumption by the network to optimize utilization, resulting in a simultaneous reduction in energy expenditures for telemedicine services.

B. *Energy-Efficient Routing Protocols*

Using energy-efficient routing protocols is essential in telemedicine networks due to the resource constraints imposed on IoT devices. Sensors and wearables typically have limited battery life and require optimization for efficient communication to forward information. Energy efficiency reduces the energy usage of 5G-enabled telemedicine networks while ensuring high levels of QoS. The resulting routing protocols aim to minimize hop counts, shorten transmission distances, or select low-energy routes while meeting the latency and bandwidth requirements of telemedicine. Table XIII summarizes fundamental routing techniques.

TABLE XII. ENERGY EFFICIENCY IN 5G NETWORK SLICING FOR TELEMEDICINE

Telemedicine application	Required network characteristics	Energy consumption (Average per session)
Real-time video consultations	Low latency and high bandwidth	High (15-20 mJ/frame)
Remote surgery assistance	Ultra-low latency and high reliability	Very high (25-30 mJ/frame)
Remote patient monitoring	Low bandwidth and moderate latency	Low (5-8 mJ/frame)

TABLE XIII. AN OVERVIEW OF ENERGY-EFFICIENT ROUTING PROTOCOLS FOR TELEMEDICINE

Routing protocol	Key feature	Energy efficiency	Use case in telemedicine
Multi-hop	Distributing data across multiple nodes	High (conserves energy for low-power devices)	Continuous patient monitoring
Opportunistic	Selecting paths based on real-time conditions	Moderate to high	Real-time video consultations
Energy-aware geographic	Using node location to select efficient routes	High	Long-term remote health monitoring with sensors

Multi-hop routing is also a common energy-conserving tactic in telemedicine networks, particularly when devices are dispersed over a large geographic area. Multi-hop routing utilizes intermediate nodes to relay data, rather than directly forwarding it to the destination node. Thus, singular IoT devices consume less energy by spreading communication burdens across multiple nodes. For instance, if a telemedicine system comprises numerous sensors that monitor patients, data from remote sensors might be relayed via short-range devices to a server. In this manner, the remote sensor is not subject to as much direct energy stress.

Optimistic routing makes accommodations according to varying network conditions, including bandwidth, node battery life, and traffic. This approach is best suited for applications with real-time requirements, such as teleconsultations, where providing QoS with minimal energy consumption is necessary. In telemedicine applications, wearable sensor devices can be replaced with high-power devices when they run out of battery. Opportunistic routing also accommodates network topology changes, such as when devices move to new locations or become temporarily disconnected.

Energy-aware geographical routing selects data transmission routes with minimum energy consumption based on the nodes' geographical positions. It is also helpful for telemedicine applications with geographically distributed units, such as rural health tracking or mobile health services. Geospatial routing tends to minimize energy usage by selecting nodes for data exchange with minimum distances. For instance, a network in a remote patient surveillance system would select a device closer to the provider for data transmission with minimal energy usage.

C. Power Control and Resource Allocation

Power control and resource allocation greatly enhance energy efficiency for telemedicine networks, particularly in 5G networks, whereby several mobile applications and devices must cooperate. As demonstrated in Table XIV, these technologies ensure that network capacities, i.e., power, bandwidth, and spectrum, are allocated to achieve performance with energy efficiency. In telemedicine applications, where continuity of service and long-term utilization of IoT devices are most significant, effective power control and resource allocation can considerably decrease energy consumption without any trade-off in QoS.

Power control systems dynamically regulate device transmission power according to real-time network conditions, energy availability, and QoS requirements. In telemedicine, where equipment such as wearable sensors and mobile diagnostic instruments has limited battery capacity, power management prolongs their operational lifespan by minimizing excessive energy usage.

Battery-aware power control and context-aware power control are the most prevalent power control technologies. Devices regulate broadcast power according to their proximity to a base station or edge server. Devices close to a base station or edge server minimize power consumption, while devices further away utilize power as needed. Devices also monitor battery levels in various ways and adjust transmission power accordingly. For instance, if battery levels hit a low point, a device minimizes power consumption to conserve power while transmitting vital data. This approach varies power depending on the use environment (e.g., essential telemedicine scenarios such as tele-surgery would focus more on power than monitor applications).

In telemedicine services, 5G networks must partition bandwidth, spectrum, and computing capabilities among multiple applications and devices. Flexible partitioning of these capabilities enables higher-priority applications, such as real-time teleconsultation via video or teleoperation, to be allocated ample resources while minimizing power consumption for lower-priority applications, such as background health tracking.

Power control and resource allocation significantly enhance energy efficiency in telemedicine networks, particularly in 5G networks, where multiple mobile applications and devices must cooperate. As demonstrated in Table XIV, these technologies ensure that network capacities, i.e., power, bandwidth, and spectrum, are allocated to achieve performance with energy efficiency. In telemedicine applications, where continuity of service and long-term utilization of IoT devices are most significant, effective power control and resource allocation can considerably decrease energy consumption without any trade-off in QoS.

TABLE XIV. POWER CONTROL AND RESOURCE ALLOCATION STRATEGIES FOR ENERGY EFFICIENCY IN 5G TELEMEDICINE NETWORKS

Category	Strategy	Description	Energy impact	Telemedicine use case
Power control	Adaptive power control	Adjusting transmission power based on distance	High (saves energy in nearby devices)	Wearable health monitoring
	Battery-aware power control	Reducing power when the battery is low	Moderate (extends device battery)	Remote diagnostic tools
	Context-aware power control	Prioritizing power based on application type	Low to high (application-dependent)	Real-time video consultations and surgery
Resource allocation	Dynamic bandwidth allocation	Modifying bandwidth based on application needs	Moderate (balances QoS and energy)	Video consultations and health monitoring
	Priority-based resource allocation	Allocating resources based on application priority	High for critical applications	Remote surgeries and emergency care
	Energy-aware resource scheduling	Scheduling tasks based on device energy levels	High energy savings	Remote monitoring and health data analytics
Hybrid	Real-time video consultations	High transmission power for low latency	High (to ensure uninterrupted service)	High bandwidth allocation
	Remote patient monitoring	Reducing transmission power	Low	Low bandwidth allocation
	Remote surgery assistance	High transmission power for reliability	Very high (critical application)	High bandwidth and processing power

Power control and resource allocation methodologies employed to enhance energy efficiency couple power control with adaptive resource allocation algorithms, allowing network resources to be utilized to their maximum while minimizing energy wastage. When network demands decrease, power levels can be reduced, along with bandwidth reallocation to lower-power sections.

V. RESULTS AND DISCUSSION

This review comprehensively evaluated existing approaches for satisfying energy-efficient video transmission in 5G-capable telemedicine systems. It was concluded that there are many converging trends and meaningful challenges between video encoding, adaptive streaming, edge computing, and network optimization. Specifically, it was concluded that employing high-efficiency video codes, such as H.265/HEVC and VP9, significantly reduces transmission energy without compromising diagnostic-grade video quality, a crucial requirement for clinical telemedicine applications. Adaptive bitrate streaming was also invaluable for balancing a trade-off between video quality and battery life under fluctuating network conditions and constrained IoT scenarios.

In addition, edge computing and caching capabilities help reduce latency while offloading processing capabilities from energy-limited IoT devices, further extending device lifespans and enhancing end-user application responsiveness for real-time applications, such as teleconsultation and telediagnostics. From a network perspective, 5G functions, including network slices and dynamic resource allocation, help provide end-users with prioritized, customized performance assurance and minimal energy usage.

To verify investigated techniques, experimental works published in the literature provide quantitative measures of codec efficiency and network behavior. For example, experimental investigations have shown that H.265/HEVC reduces the bitrate by up to 50% compared to H.264/AVC for perceptually equal video quality, with a concomitant reduction of up to 40% in energy consumption in battery-powered devices streaming 1080p streams at 30 frames per second over limited networks. VP9 demonstrates similar compression efficiency, though with slightly higher computational complexity, resulting in trade-offs between encoding latency and device power usage. Simulation studies using telemedicine scenarios further show that H.265 achieves a 25–30% reduction in total energy consumption per frame when applied with adaptive bitrate streaming strategies on wearable diagnostic devices.

Additionally, trials for 5G network-slicing-based remote-surgery applications demonstrated end-to-end latencies repeatedly under 10 milliseconds with an energy-aware edge- and core network scheduling. In these trials of remote-surgery applications in managed hospital-to-hospital networks with URLLC slices, there exists not only ultra-low latency potential but also a device-side reduction in energy consumption by a factor of 20 through localized video pre-processing. Parameters like these provide robust real-world validation for feasibility and the promise of these technologies for life-critical clinical applications.

With further advancements in telemedicine, telemedicine systems will require energy-aware transmission technologies for 5G networks and IoT applications. Future research in this area could extend along a range of avenues to enhance the effectiveness, scalability, and energy efficiency of telemedicine systems, while ensuring the security and accessibility of healthcare services. The following research avenues will facilitate this field of study.

Utilizing Artificial Intelligence (AI) and machine learning is one potential area for study to enhance video streaming and encoding. ABR streaming supports active decisions in real-time based on network conditions, device battery levels, and user behavior, with a focus on balancing the trade-off between energy usage and video quality. AI models can examine network utilization patterns and device usage patterns to forecast bandwidth changes or battery consumption, enabling the active modification of encoding parameters. AI-focused solutions can significantly improve energy efficiency for telemedicine applications that are severely limited in resources, such as wearable health sensors or transportable diagnostic equipment.

Future studies should further examine the integration of telemedicine with edge computing for video transmission. Edge server-based computation of video data near IoT devices reduces the need to transport high-definition video over long distances, thereby conserving bandwidth and power. Research could also utilize video compression, edge-level transcoding, and real-time processing to alleviate computational burdens on resource-poor devices. This would enable battery-powered mobile health units to support high-definition video consultations via telemedicine. Furthermore, integrating technology between edge computing and 5G ultra-reliable low-latency communications may further enhance telemedicine for long-distance interactions.

Blockchain technology can potentially enhance resource allocation in 5G telemedicine networks. The decentralized ledger of blockchain guarantees transparent, equitable, and efficient allocation of bandwidth and power resources across various telemedicine applications. Future research may focus on designing blockchain-based smart contracts that automate the distribution of network resources for video transmission, taking into account device energy levels, network congestion, and QoS requirements. This approach would allow a fairer and more energy-efficient use of 5G infrastructure in telemedicine settings, particularly when managing multiple simultaneous users and applications.

Despite existing compression standards, such as H.265/HEVC and VP9, that provide significant video efficiency improvements, further research on future-generation video compression technologies is needed for telemedicine. Quantum computing-based video compression and neural network-controlled codec-based compression may offer larger compression ratios without compromising video quality, resulting in significantly reduced bandwidth and energy usage. This technology will be most effective for telemedicine applications that demand the continuous transmission of high-resolution video images, such as tele-operations and tele-diagnostics, through real-time video.

Investigating battery-conscious transmission methods for videos can promote energy efficiency by enabling telemedicine networks to adjust bitrate and video quality while preserving battery life in IoT devices. Corresponding protocols would allow IoT devices to consume less power during non-essential times or with depleted batteries while allocating more power for vital occasions or emergencies. Integrating power reduction algorithms in future telemedicine networks can promote wearable device and sensor lifespan while minimizing recharging or battery replacement.

Forthcoming research must prepare for 6G and subsequent technologies. 6G will prioritize green communications, highlighting sustainability and energy efficiency. Researchers can also study the implementations of 6G's ultra-low power transmission techniques, quantum telecommunication, and terahertz radiation in telemedicine networks to further reduce energy usage. Reviews of early work on integrating 6G technology with telemedicine services promise to steer future innovation toward energy-aware healthcare provision.

Machine-learning advancements permit us to achieve predictive ABR streaming, which forecasts bandwidth changes before they occur and provides adaptive adjustments in encoding parameters. Time-series network throughput predictions benefit most from deep learning models, such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). LSTM-based ABR controllers, for example, achieve up to 90% accuracy in future bandwidth-window predictions, allowing devices to prefetch video segments with optimal bitrate pre-buffering, resulting in reduced re-encoding energy consumption and playback stalling. They can be trained from historical network data and installed on edge nodes or IoT gateways with minimal computation overhead.

VI. Conclusion

5G technology and energy-constrained IoT devices have great potential to convert telemedicine into telemedicine services with real-time, high-grade services. However, one of the most significant concerns is the deployment of energy-aware video transmission, ensuring that IoT devices remain active for extended periods without compromising QoS. This paper presents a detailed review of various energy-aware methodologies, including video encoding schemes, adaptive bitrate streaming, edge computing, caching, network slicing, and power control policies, which play a critical role in reducing energy consumption for telemedicine applications.

By employing sophisticated compression technologies such as H.265/HEVC for videos, along with adaptive bitrate streaming and edge computing, telemedicine deployments can effectively reduce power usage for transmission while maintaining the high definition necessary for accurate diagnosis and consultations. 5G's network slice capabilities also provide for flexible network resource allocation so that mission-critical applications such as tele-surgeries receive whatever bandwidth and low-latency services they require while day-to-day monitoring operations consume little power.

Power control and resource allocation methodologies, such as adaptive power control and battery-aware routing, improve the energy efficiency of telemedicine networks. These methodologies enable IoT devices to communicate with reduced energy consumption levels, thereby extending device lifespan with minimal frequent recharging. Furthermore, future advancements in AI-based video encoding, blockchain-based resource allocation, and emerging telecommunication technologies such as 6G offer further potential for improvements in telemedicine energy efficiency.

References

[1] G. Sageena, M. Sharma, and A. Kapur, "Evolution of smart healthcare: Telemedicine during COVID-19 pandemic," Journal of The Institution of Engineers (India): Series B, pp. 1-6, 2021.

[2] B. Pourghebleh and V. Hayyolalam, "A comprehensive and systematic review of the load balancing mechanisms in the Internet of Things," Cluster Computing, pp. 1-21, 2019.

[3] D. Kumari, S. Sharma, M. Chawla, and S. Panda, "A Manifesto for Healthcare Based Blockchain: Research Directions for the Future Generation," Journal of The Institution of Engineers (India): Series B, pp. 1-22, 2024.

[4] E. H. Golembiewski et al., "Video-based observation research: A systematic review of studies in outpatient health care settings," Patient Education and Counseling, vol. 106, pp. 42-67, 2023.

[5] F. Kamalov, B. Pourghebleh, M. Gheisari, Y. Liu, and S. Moussa, "Internet of medical things privacy and security: Challenges, solutions, and future trends from a new perspective," Sustainability, vol. 15, no. 4, p. 3317, 2023.

[6] M. M. Nasralla, S. B. A. Khattak, I. Ur Rehman, and M. Iqbal, "Exploring the role of 6G technology in enhancing quality of experience for m-health multimedia applications: A comprehensive survey," Sensors, vol. 23, no. 13, p. 5882, 2023.

[7] S. A. Butt, A. Ali, A. Khalid, Q. Shaheen, P. Ariza-Colpas, and M. Shiraz, "Patients using real-time remote health monitoring applications: a review," Computational Intelligence for Medical Internet of Things (MIoT) Applications, pp. 93-120, 2023.

[8] N. Pegoraro et al., "Telemedicine in Sports under Extreme Conditions: Data Transmission, Remote Medical Consultations, and Diagnostic Imaging," International journal of environmental research and public health, vol. 20, no. 14, p. 6371, 2023.

[9] K. C. Maita et al., "Imaging evaluated remotely through telemedicine as a reliable alternative for accurate diagnosis: a systematic review," Health and Technology, vol. 13, no. 3, pp. 347-364, 2023.

[10] K. Khujamatov, N. Akhmedov, E. Reypnazarov, D. Khasanov, and A. Lazarev, "Device-to-device and millimeter waves communication for 5G healthcare informatics," in Blockchain Applications for Healthcare Informatics: Elsevier, 2022, pp. 181-211.

[11] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, "5G technology for healthcare: Features, serviceable pillars, and applications," Intelligent Pharmacy, vol. 1, no. 1, pp. 2-10, 2023.

[12] A. M. Peralta-Ochoa, P. A. Chaca-Asmal, L. F. Guerrero-Vásquez, J. O. Ordoñez-Ordoñez, and E. J. Coronel-González, "Smart healthcare applications over 5G networks: A systematic review," Applied Sciences, vol. 13, no. 3, p. 1469, 2023.

[13] A. Sacco, F. Esposito, G. Marchetto, G. Kolar, and K. Schwetye, "On edge computing for remote pathology consultations and computations," IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 9, pp. 2523-2534, 2020.

[14] D. K. Israni and N. S. Chawla, "Human - Machine Interaction in Leveraging the Concept of Telemedicine," Human - Machine Interface: Making Healthcare Digital, pp. 211-245, 2023.

[15] B. Pourghebleh, K. Wakil, and N. J. Navimipour, "A comprehensive study on the trust management techniques in the Internet of Things," IEEE Internet of Things Journal, vol. 6, no. 6, pp. 9326-9337, 2019.

[16] B. Pourghebleh, N. Hekmati, Z. Davoudnia, and M. Sadeghi, "A roadmap towards energy - efficient data fusion methods in the Internet of Things," Concurrency and Computation: Practice and Experience, vol. 34, no. 15, p. e6959, 2022.

[17] S. Islam, A. K. Budati, M. K. Hasan, S. Goyal, and A. Khanna, "Performance analysis of video data transmission for telemedicine applications with 5G enabled Internet of Things," Computers and Electrical Engineering, vol. 108, p. 108712, 2023.

[18] S. Bouaafia, R. Khemiri, S. Messaoud, O. Ben Ahmed, and F. E. Sayadi, "Deep learning-based video quality enhancement for the new versatile video coding," Neural Computing and Applications, vol. 34, no. 17, pp. 14135-14149, 2022.

[19] Z.-N. Li, M. S. Drew, J. Liu, Z.-N. Li, M. S. Drew, and J. Liu, "Modern video coding standards: H. 264, H. 265, and H. 266," Fundamentals of Multimedia, pp. 423-478, 2021.

[20] O. Nawaz, M. Fiedler, and S. Khatibi, "QoE-Based Performance Comparison of AVC, HEVC, and VP9 on Mobile Devices with Additional Influencing Factors," Electronics, vol. 13, no. 2, p. 329, 2024.

[21] A. Rana, C. Chakraborty, S. Sharma, S. Dhawan, S. K. Pani, and I. Ashraf, "Internet of medical things-based secure and energy-efficient framework for health care," Big Data, vol. 10, no. 1, pp. 18-33, 2022.

[22] D. Zhang, L. Wei, K. Shen, H. Zhu, D. Wang, and F. Wang, "TrimStream: Adaptive Real-time Video Streaming through Intelligent Frame Retrospection in Adverse Network Conditions," IEEE Transactions on Mobile Computing, 2024.

[23] T. Kimura, T. Kimura, A. Matsumoto, and K. Yamagishi, "Balancing quality of experience and traffic volume in adaptive bitrate streaming," IEEE Access, vol. 9, pp. 15530-15547, 2021.

[24] D. Alekseeva, A. Ometov, O. Arponen, and E. S. Lohan, "The future of computing paradigms for medical and emergency applications," Computer Science Review, vol. 45, p. 100494, 2022.

[25] J. Lin, L. Zhu, W.-M. Chen, W.-C. Wang, and S. Han, "Tiny machine learning: Progress and futures [feature]," IEEE Circuits and Systems Magazine, vol. 23, no. 3, pp. 8-34, 2023.

[26] S. Wijethilaka and M. Liyanage, "Survey on network slicing for Internet of Things realization in 5G networks," IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 957-994, 2021.