Effective hybrid video denoising and blending framework for Internet of Remote Things (IoRT) environments

B. Aruna Devi & Mani Deepak Choudhry

Published online: 02 Feb 2024.

Submit your article to this journal ⌴

Article views: 519

View related articles ⌴

View Crossmark data ⌴

Citing articles: 1 View citing articles ⌴

OPEN ACCESS

Check for updates

Effective hybrid video denoising and blending framework for Internet of Remote Things (IoRT) environments

B. Aruna Devi[a] and Mani Deepak Choudhry[b]

[a]Department of Electronics and Communication Engineering, Dr. N.G.P. Institute of Technology, Coimbatore, India; [b]Department of Information Technology, KGiSL Institute of Technology, Coimbatore, India

ABSTRACT

The Internet of Remote Things (IoRT) has emerged as a transformative paradigm, merging IoT capabilities with remote technologies. IoRT environments, featuring interconnected sensors and robots, face challenges like sensor noise and low-light conditions, compromising video stream quality. This paper proposes a Hybrid Video Denoising and Blending Framework to address IoRT video data shortcomings. Leveraging spatial and temporal domain denoising techniques, the framework effectively removes noise while preserving crucial details. The inclusion of advanced blending algorithms facilitates seamless fusion of data from multiple sources, enhancing decision-making in real-world scenarios. The framework adopts a dynamic weighted averaging approach and an optimal sensor selection mechanism to intelligently choose informative data sources, improving blended output quality. Extensive experiments with a diverse IoRT dataset showcase the framework's superiority over state-of-the-art techniques, offering significant enhancements in video quality, noise reduction, and data fusion accuracy. Applications like surveillance, autonomous remotes, and industrial automation can benefit from the framework's ability to provide clearer, more reliable visual information. In conclusion, this research introduces a pioneering approach to mitigate video noise and enhance data fusion in IoRT, showcasing promising results and paving the way for further research in the integration of Remotes and IoT.

1. Introduction

In an era where interconnected devices are reshaping industries and our daily lives, the Internet of Things (IoT) has emerged as a transformative force. IoT technologies have extended their influence beyond urban landscapes, finding applications in the most remote and challenging environments. This expansion has given rise to the concept of the Internet of Remote Things (IoRT), a subset of IoT tailored to environments where direct human interaction is limited, dangerous, or even impossible. IoRT brings with it a new set of possibilities and complexities, particularly when it comes to processing video data a cornerstone of modern information exchange.

1.1. The unveiling of IoRT and its expansive implications

The term "Internet of Remote Things (IoRT)" refers to the integration of remote or distant devices, sensors, and systems into the broader framework of the Internet of Things (IoT). While the traditional IoT focuses on connecting devices within a localized environment, IoRT extends this concept to include objects and technologies that operate in remote or distant locations. The key distinguishing factor is the geographical separation of the devices, which can be situated in areas that are challenging to access or are not part of the immediate physical vicinity.

The concept of the Internet of Remote Things (IoRT) gains significance as our technological landscape evolves, pushing the boundaries of what is achievable with connectivity. In the realm of healthcare, IoRT can be envisioned through the deployment of medical sensors and devices in remote patient monitoring. Patients residing in distant locations or those with limited access to healthcare facilities can benefit from continuous monitoring of vital signs, enabling healthcare professionals to remotely assess and respond to their health conditions in real time. This not only enhances patient care but also contributes to the early detection of health issues.

IoRT is not limited to terrestrial applications; it extends into the exploration of space. In space missions, where human presence is limited or non-existent, IoRT can be instrumental in gathering data and controlling devices on distant planets or celestial bodies. Robotic explorers equipped with IoRT capabilities can navigate and perform tasks in extraterrestrial environments, relaying crucial information back to Earth for analy-

---

CONTACT Mani Deepak Choudhry ✉ manideepak.c@kgkite.ac.in

sis and decision-making. The energy sector is another domain where IoRT can revolutionize operations. In remote areas with energy infrastructure, such as offshore wind farms or isolated power stations, IoRT enables efficient monitoring and management of equipment. Sensors and smart devices can detect anomalies, optimize energy production, and enhance maintenance procedures, ultimately ensuring the reliability and sustainability of energy sources in challenging environments. One notable aspect of IoRT is its ability to bridge geographical gaps and overcome logistical challenges. By incorporating remote devices into the IoT ecosystem, industries can enhance efficiency, optimize resource utilization, and gather valuable insights from previously inaccessible locations. For instance, in agriculture, IoRT might involve deploying sensors in remote fields to monitor soil conditions, weather patterns, and crop health, enabling farmers to make data-driven decisions even in distant agricultural landscapes.

In essence, the Internet of Remote Things extends the reach and impact of the IoT paradigm, bringing connectivity and intelligence to remote environments where traditional connectivity solutions might be impractical or unfeasible. As technology continues to advance, IoRT holds the promise of unlocking new possibilities for remote monitoring, automation, and control across a diverse range of industries and applications. As IoRT continues to evolve, it brings forth a paradigm shift in how we perceive and interact with technology across vast distances. The ability to connect and control devices in remote locations not only opens new avenues for exploration and industry but also fosters a more interconnected and intelligent world, where the benefits of technology can reach even the most distant corners of our planet and beyond.

1.2. Navigating challenges in video data quality within IoRT

In the Internet of Robotic Things (IoRT) environments, several challenges pose significant hurdles to the seamless operation and advancement of robotic technologies. One such challenge is the constraint on bandwidth. The interconnected nature of robotic devices requires a substantial amount of data exchange for real-time communication and coordination. Limited bandwidth can lead to delays in data transmission, impacting the responsiveness of robotic systems. This constraint is particularly critical in scenarios where split-second decisions are vital, such as in autonomous vehicles or emergency response robots. The implication is a potential compromise in the efficiency and reliability of IoRT applications, potentially hindering their widespread adoption in time-sensitive domains.

In the realm of the Internet of Robotic Things (IoRT), challenges abound, with noisy data posing a significant hurdle due to environmental factors such as interference and sensor malfunctions that can distort crucial sensor information. Overcoming these challenges is essential to ensure the reliability and safety of robotic technologies. Furthermore, security concerns escalate with the increased integration of robots across various industries, where cyber-attacks and unauthorized access could compromise robotic systems, especially in critical applications like healthcare. Interoperability issues also emerge as a hurdle in IoRT, as the lack of standardized communication protocols may lead to fragmentation, limiting the scalability and adaptability of robotic technologies across diverse applications. In parallel, real-time responsiveness is critical in IoRT scenarios, where any latency or video quality degradation risks delays and system efficiency. Managing data transmission and bandwidth becomes challenging in interconnected environments, necessitating efficient compression algorithms and networking solutions. High-quality video streams demand substantial bandwidth, a key factor for secure transmission and storage, particularly in human-robot interaction applications. The clarity of visuals in Human-Robot Interaction (HRI) becomes pivotal for effective communication, underscoring the significance of video quality in seamlessly integrating robotic technologies. Addressing these multifaceted challenges is imperative for unlocking the full potential of IoRT devices and fostering enhanced safety, efficiency, and user acceptance.

1.3. Paving the way for advanced video processing in IoRT

Denoising techniques play a crucial role in enhancing the quality of images by reducing unwanted noise, and several algorithms have been developed to address this challenge.

- Non-local Means (NLMeans) is a widely used denoising algorithm that excels in preserving image details by averaging similar patches from different regions. A
- Another powerful denoising algorithm is Block-Matching 3D (BM3D), which employs collaborative filtering to reduce noise in 3D groups of similar blocks. BM3D has proven successful in scenarios where noise occurs in complex patterns, such as low-light conditions or images affected by compression artifacts.
- Wavelet denoising leverages wavelet transforms to decompose images into different frequency components, offering a versatile approach to noise removal across various scales. This technique is particularly useful in applications where noise manifests at different frequencies or when a multi-scale analysis is essential.
- Total Variation Denoising minimizes the total variation of an image, making it well-suited for situations

where preservation of fine details is paramount. Medical imaging applications, for instance, benefit from this technique as it effectively reduces noise while maintaining the integrity of intricate structures.

On the blending front, techniques such as Alpha Blending provide a fundamental approach to combining images, assigning weights to pixels to control the intensity of each image's contribution. Widely used in image overlays and compositing, Alpha Blending is efficient and straightforward. While its simplicity is an advantage, performance evaluations highlight potential challenges in achieving seamless transitions, particularly in scenarios where image characteristics significantly differ.

Poisson Blending focuses on achieving seamless transitions between images through solving a Poisson equation, essential for image stitching and panorama creation. Multi-resolution Blending, blending images at different resolutions, is valuable in virtual reality and HDR imaging, ensuring efficient and visually appealing results. Exposure Fusion excels in HDR photography, blending multiple exposures for an extended dynamic range. Gradient Domain Blending minimizes gradient differences for seamless transitions in image stitching and compositing. Traditional video processing methods fall short in IoRT, prompting the proposal of a hybrid framework designed to tackle challenges like noise and quality degradation. The forthcoming sections will delve into the methodology, experiments, and results showcasing the potential of this hybrid framework to revolutionize IoRT video data processing.

2. Related works

This survey's primary aim is to explore the domain of video processing within IoRT environments, underlining the necessity to enhance video quality for better functionality and performance of remote systems. As IoRT continues to burgeon, comprehending and ameliorating video processing techniques become essential. Several key research questions will steer this study, including inquiries into the optimization of video denoising and blending techniques for IoRT, challenges and solutions in video processing in remote IoRT settings, and the impact of different video processing techniques on bandwidth and storage in IoRT applications.

2.1. Background and concepts

IoRT is a sophisticated framework that extends the connectivity capabilities of the Internet of Things (IoT) to devices and systems in remote, often isolated locations, enabling them to transmit and receive data for monitoring, control, and automation purposes [1]. This

technology is particularly crucial in areas where traditional network infrastructure is either non-existent or impractical to deploy.

The architecture of IoRT is a complex, multi-layered system comprising several key components:

(1) **Sensors/Actuators**: At the foundation of the IoRT are remote sensors and actuators. Sensors collect various types of data from the environment, such as temperature, pressure, or images, while actuators perform actions based on the processed data, like adjusting a thermostat or activating a pump.
(2) **Connectivity**: This layer involves the communication networks that connect these remote devices to the internet.
(3) **Data Processing**: Data transmitted by sensors is processed either at the edge (near or on the device) or in the cloud.
(4) **User Interface**: The top layer is where humans interact with the IoRT system, often through dashboards, mobile apps, or web applications, allowing users to monitor data, receive alerts, and perform manual overrides [2–6].

In the realm of video processing, techniques like video denoising and blending are essential in IoRT. Video denoising [7] is the process of removing noise or graininess from video footage, which is common in data transmitted over long distances or through suboptimal networks. This process helps in enhancing the clarity and overall quality of the video [8–10]. Video blending involves combining video data from multiple sources or frames to produce a single, high-quality output [11,12].

However, video processing in IoRT presents significant challenges. Limited bandwidth means that transmitting high-quality video is difficult, often necessitating local processing and smart compression techniques. High latency and intermittent connectivity can delay the transmission of crucial real-time video data [13], making quick decision-making challenging. Furthermore, the need for real-time processing in potentially unstable environments requires robust [14], fault-tolerant systems and advanced algorithms that can make intelligent decisions locally [15]. Addressing these challenges necessitates innovative solutions in video processing technology, network infrastructure, and data analytics, making this an exciting and rapidly evolving field of research and development

2.2. Video denoising techniques

Traditional video denoising techniques primarily revolve around temporal and spatial filtering.

(1) **Temporal Filtering**: This technique leverages the information from successive frames in a video sequence. By analyzing the differences and

similarities across these frames, temporal filtering aims to reduce noise that varies between frames while preserving the actual motion and details within the scene [16]. However, in the context of IoRT, the effectiveness of temporal filtering can be limited due to high latency in data transmission, which disrupts the sequence of real-time video frames [17].

(2) **Spatial Filtering**: Spatial filtering, on the other hand, focuses on reducing noise within a single frame rather than across a sequence [18]. It involves techniques like Gaussian blurring or median filtering, which work by analyzing the pixels around a target pixel and recalculating the target pixel's value based on its neighbours [19]. While this can be effective for static noise, it often results in a loss of detail, leading to blurred images. In IoRT, the challenge intensifies due to the variable quality of transmitted images and the need for real-time processing, often leading to either over-smoothed or still noisy results [20].

Deep learning models, especially Convolutional Neural Networks (CNNs) [21] and autoencoders, have shown significant promise in distinguishing between true image content and noise [22], providing superior denoising results with better detail preservation. ML and DL offer substantial improvements in video denoising quality, their practical application in IoRT is contingent on the availability of suitable computational resources and the specific requirements of the IoRT application in question. As technology advances, we anticipate more efficient models and edge computing solutions that could make these advanced denoising techniques more accessible and suitable for IoRT environments [23–26].

2.3. Video blending techniques

To seamlessly merge video streams in the Internet of Remote Things (IoRT), essential blending techniques include alpha blending, which combines videos using transparency factors for smooth transitions but may lead to ghosting or double exposures with diverse or fast-moving scenes. Pyramid blending, a more advanced method, breaks down images into frequency layers and blends them sequentially, offering a sophisticated approach to creating cohesive video streams [27]. These methods play a crucial role in ensuring fluid and artifact-free transitions for IoRT applications. Maintaining consistency across video streams is complex, given the potential for varying environmental conditions and camera settings [28]. Moreover, ensuring quality is a significant hurdle, as network issues like limited bandwidth and high latency are common in IoRT scenarios, potentially disrupting the real-time transmission and processing of video data [29]. In IoRT, addressing constraints involves innovative solutions such as utilizing edge computing to process video data near the source, reducing latency and bandwidth usage. Additionally, advanced algorithms can dynamically adjust blending parameters in response to network conditions and video content in real-time. Leveraging machine learning models to predict and compensate for network delays offers the potential for smoother video streams in IoRT applications, enhancing the efficiency of remote operations [29,30].

2.4. Hybrid frameworks for video processing

In the context of the Internet of Robotic Things (IoRT), a progressive approach involves combining video denoising and blending to enhance the quality of video data, especially in remote environments with unstable networks. Emerging frameworks integrate both processes, offering sequential, parallel, or integrated processing approaches, with sequential processing being simpler but potentially challenging in real-time applications due to latency [31–35]. Simultaneously, the integration of AI, particularly deep learning, is revolutionizing video processing frameworks by training models to identify and remove noise while seamlessly blending video streams in a unified approach [36]. When coupled with edge computing in IoRT, these AI models enable more efficient video processing near the data source, effectively reducing latency. This synergy of AI and edge computing holds significant promise for enhancing video quality, thereby making remote monitoring and automation systems in IoRT more reliable, efficient, and high-quality.

2.5. Challenges and solutions in IoRT video processing

The Internet of Remote Things (IoRT) brings forth a unique set of challenges, primarily due to the remote environments in which it operates, often characterized by connectivity issues, limited power, scarce computational resources, and heightened concerns around data security and privacy [37–39].

(1) **Connectivity Fluctuations**: Remote areas often suffer from unstable internet connections or, in some cases, rely on satellite communication, which can be both slow and expensive.

(2) **Limited Power and Computational Resources**: Devices in remote locations often have restricted access to power, relying on batteries or renewable sources. Additionally, the computational capacity of these devices is often limited, making it challenging to process complex algorithms locally.

(3) **Data Security**: The transmission of data over long distances, potentially over unsecured or public networks, increases the risk of interception or unauthorized access.

(4) **Privacy Concerns**: Many IoRT applications collect sensitive information. Ensuring this data is handled and stored securely is paramount to maintaining user trust and compliance with privacy laws.

Existing solutions have sought to address these challenges with varying degrees of success:

- **Edge Computing**: By processing data closer to where it is generated, edge computing addresses several of these issues. It reduces the need for constant connectivity and the amount of data that needs to be transmitted, conserving bandwidth.
- **Data Encryption and Secure Protocols**: The use of end-to-end encryption and secure communication protocols like TLS/SSL can significantly enhance data security. However, these methods can also increase the computational load and require a stable connection to maintain a continuous security handshake between devices.
- **Privacy-Preserving Algorithms**: Techniques like federated learning and differential privacy enable analysis without requiring access to raw data, helping mitigate privacy concerns. However, they can be complex to implement and may not be suitable for all types of analysis.
- **Energy-Efficient Hardware and Algorithms**: The development of low-power hardware and energy-efficient algorithms has been crucial for power-constrained IoRT devices. However, there's often a trade-off between power efficiency and computational capacity.

Recent advances in video denoising, blending, and the Internet of Robotic Things (IoRT) are driven by breakthroughs in artificial intelligence, computer vision, and robotics. Deep learning, utilizing convolutional and recurrent neural networks, enhances video denoising, preserving fine details while eliminating noise effectively. In blending technologies, applications from augmented reality to video editing benefit from deep learning approaches, such as generative adversarial networks, improving the quality of blended content across various industries like entertainment and virtual collaboration. In IoRT, the integration of robotics and the internet facilitates seamless collaboration among robots, enhancing efficiency across industries. Real-time data exchange allows robotic systems to adapt to dynamic environments, fostering the development of smart factories and autonomous vehicles. The convergence of video processing and IoRT is promising, as integrating denoising and blending into robotic vision enhances perception, enabling informed decision-making and improved interaction. Considering scalability, solutions must be tailored to the specific requirements of each IoRT application and environment.

3. Proposed scheme

The framework combines adaptive thresholding and Fourier Transform-based filtering for denoising and employs a weighted average approach for blending, optimizing the visual quality, energy consumption, and latency in various IoRT settings. In IoRT environments, the transmission of high-quality video is crucial for various applications, including surveillance, remote monitoring, and telemedicine. The challenges posed by noisy environments, limited bandwidth, and resource constraints necessitate the development of sophisticated video processing techniques. The proposed framework addresses these challenges by integrating advanced denoising and blending methodologies, ensuring seamless video transmission with minimal resource utilization and latency.

In manufacturing and Industry 4.0, a hybrid framework seamlessly integrates traditional and collaborative robots, utilizing AI algorithms for predictive maintenance and quality control with real-time sensor data. For autonomous vehicles, the hybrid system combines edge computing for immediate decision-making by onboard AI with cloud-based analytics for long-term traffic pattern identification and route optimization. In healthcare, the hybrid approach integrates robotic assistants using local processing for immediate patient interaction and cloud-based AI for complex diagnostics, enhancing patient care. In smart agriculture, the hybrid framework optimizes precision farming through real-time data processing on drones and ground-based robots, coupled with cloud-based analysis for sustainable practices. Warehouse operations benefit from the hybrid framework, combining AGVs and robotic arms with on-board AI for immediate navigation and cloud-based analytics for long-term efficiency improvements in inventory management and order fulfilment.

3.1. Algorithm

Input:

- Video Frames: $V\_in(x, y, t)$
- Blending Factor: α
- Weights for each frame: w_i.
- Gaussian Filter Standard Deviation: σ
- Noise Standard Deviation: σ_n.
- Window Size for Adaptive Thresholding: N.
- Constants: β, γ
- Number of Hops in Network: M.
- Propagation Time for each Hop: T_p_i.
- Amount of Data Transmitted in each Hop: D_i.
- Data Rate for each Hop: R_i.
- Constants for Latency Calculation: δ, ϵ
- Transmission Power: P_tx.
- Gains of Transmitting and Receiving Antennas: G_tx, G_rx.

- Wavelength: λ
- Distance between Transmitter and Receiver: d.
- System Loss: L.
- Weights for QoE: $\omega\_1, \omega\_2, \omega\_3$.
- Threshold Values for PSNR and SSIM: θ_1, θ_2.
- Standard Deviations for PSNR and SSIM: σ_1, σ_2.

Output:

- Denoised and Blended Video Frame: V_out(x, y, t)
- Total Latency: L_total.
- Total Energy Consumption: E_total.
- Quality of Experience: Q.

Algorithm Steps:

(1) For each video frame V in(x, y, t) do
  (a) Compute Adaptive Threshold T(x, y, t) using

$$T(x, y, t)$$
$$= \beta * (1/N * \Sigma(\Sigma V\_in(x + i, y + j, t)))$$
$$+ \gamma * sqrt(1/N^2 * \Sigma(\Sigma(V\_in(x$$
$$+ i, y + j, t))^2))$$

  (b) Apply Fourier Transform-based filtering and compute V_out(x, y, t) using

$$V\_out(x, y, t)$$
$$= \alpha * (\Sigma w_i * F^{(-1)}[(F[V\_in(x, y, t)]$$
$$* e^{(-((x-\mu_x)^2 + (y-\mu_y)^2)/(2\sigma^2))})/$$
$$(1 + (\sigma_n/F[V\_in(x, y, t)])^2)])$$
$$+ (1 - \alpha) * V\_blend(x, y, t)$$

(2) Compute V_blend(x, y, t) using

$$V\_blend(x, y, t) = \Sigma w_k * V_k(x, y, t)/\Sigma w_k$$

(3) Compute Total Latency L_total using

$$L\_total = \Sigma(T_{p_i} + D_i/R_i)$$
$$+ \delta * log2(1 + P_t x * G_t x * G_r x * \lambda^2/$$
$$(4\pi)^2 * d^2 * L) + \varepsilon * H(X)$$

(4) Compute Total Energy Consumption E_total using

$$E\_total = \int P\_active(t)dt + P\_idle(T - t\_active)$$

(5) Compute Quality of Experience Q using

$$Q = \omega_1 * (1 - e^{(-(PSNR-\theta_1)/\sigma_1)})$$
$$+ \omega_2 * (1 - e^{(-(SSIM-\theta_2)/\sigma_2)})$$
$$- \omega_3 * L\_total$$

(6) Return $V\_out(x, y, t), L\_total, E\_total,$ and Q

3.1.1. *Effective hybrid video denoising and blending framework*

The framework employs a hybrid denoising approach that combines adaptive thresholding and Fourier Transform-based filtering. The denoising process is mathematically represented by a comprehensive equation, considering various parameters like blending factor, frame weights, and standard deviation of the Gaussian filter. Adaptive thresholding is meticulously calculated to optimize the denoising process, considering the local characteristics of each video frame. The methodology amalgamates advanced denoising and blending techniques, aiming to elevate the quality of experience, optimize energy consumption, and minimizes data transmission latency in diverse IoRT settings as shown in Figure 1.

(i) Median Filter for Denoising

The median filter algorithm is a non-linear digital filtering technique used primarily for noise reduction in images and videos. This process involves iterating over each pixel in an image or a frame of a video, then, for each pixel, examining a surrounding window of neighbouring pixels – the size of which is defined by a predefined windowSize (e.g. 3×3, 5×5). The pixel values in this window are sorted numerically, and the median value (the middle pixel in the sorted list) is computed. The original pixel is then replaced with this median value in the output image or frame, effectively reducing noise while preserving edges within the image.

(ii) Cuckoo Search for Optimization

The Cuckoo Search Optimization algorithm for video tailors a nature-inspired method based on cuckoo bird behaviour to optimize video parameters. Initially, a population of "nests" representing potential video processing parameters is generated. As iterations progress, each "nest" undergoes adjustments based on Levy flights to find better parameter combinations. A random nest can be replaced if a new combination proves superior. To introduce randomness and escape local optima, with a certain probability, the worst-performing nests are abandoned and replaced with new random parameter sets. The algorithm evaluates each nest's quality by applying its parameters to the video and measuring specific metrics, like clarity or compression efficiency. The process iterates until predefined conditions (like a maximum number of generations) are met, ultimately returning the best video parameters discovered.

(iii) Proposed Hybrid Maximum likelihood estimation (MLE) – Maximum dynamic range (MDR) for blending

The Hybrid MLE-MDR Blending algorithm combines the statistical robustness of Maximum Likelihood Estimation (MLE) with the dynamic range
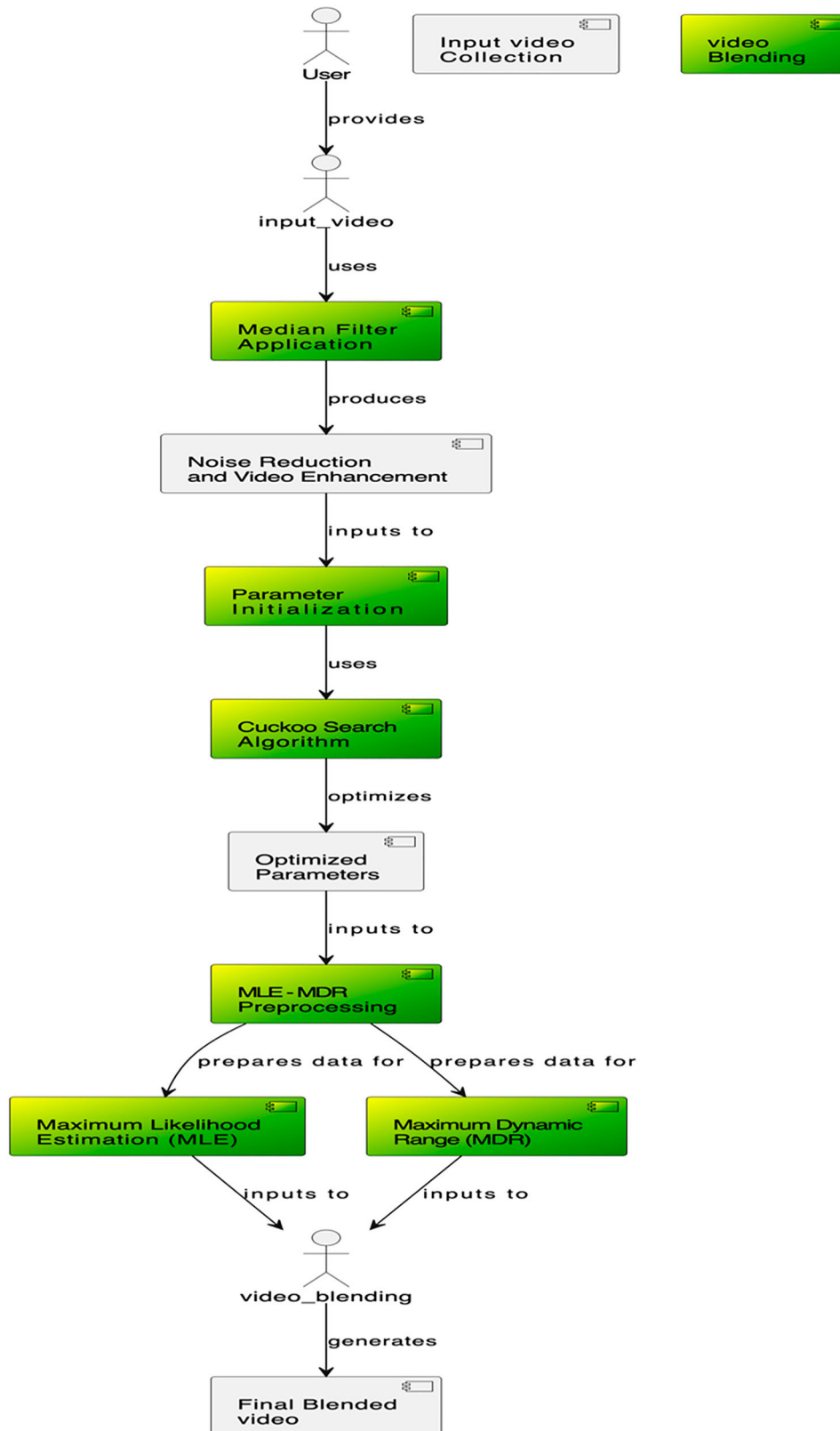
Figure 1. Architecture Diagram.

enhancement of Maximum Dynamic Range (MDR) to blend multiple images. For each pixel location across the input images, the algorithm calculates a weighted mean based on a specified weighting function and the noise level (sigma). This mean is then adjusted using the MDR method, which factors in the difference between the highest and lowest pixel values from the input images, ensuring the final blended image retains good contrast and brightness. The blended pixel value is a mix of these MLE and MDR calculations, resulting in a harmonized image that incorporates the best attributes from the input images.

3.2. MLE and MDR based video frame blending and cuckoo for optimization

The framework is rigorously evaluated in a simulated IoRT environment, comparing the Quality of Experience (QoE), energy consumption, and latency with existing methodologies. The optimization process involves fine-tuning various parameters to achieve the optimal trade-off between video quality, energy consumption, and latency, aligning with the specific requirements and constraints of IoRT environments. This approach is meticulously represented by the equation:

$$V_{\text{out}}(x, y, t)$$
$$= \alpha \left(\sum_{i=1}^{N} w_i \mathcal{F}^{-1} \left[\frac{\mathcal{F}[V_{\text{in}}(x, y, t)] \cdot e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}}}{1 + \left(\frac{\sigma_n}{\mathcal{F}[V_{\text{in}}(x,y,t)]} \right)^2} \right] \right)$$
$$+ (1 - \alpha) V_{\text{blend}}(x, y, t) \tag{1}$$

This equation is instrumental in reducing noise in video frames, considering various parameters like the blending factor, α, the number of frames being averaged, N, the weights for each frame, wi, and the standard deviation of the Gaussian filter, σ, among others. Additionally, the adaptive thresholding is calculated using the equation:

$$T(x, y, t)$$
$$= \beta \left(\frac{1}{N} \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} V_{\text{in}}(x + i, y + j, t) \right)$$
$$+ \gamma \sqrt{\frac{1}{N^2} \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} (V_{\text{in}}(x + i, y + j, t))^2} \tag{2}$$

This equation optimizes the denoising process by considering the local characteristics of each video frame, utilizing parameters like the window size, N, and constants, β and γ. For blending video frames, a weighted average methodology is adopted, represented by the equation:

$$V_{\text{blend}}(x, y, t) = \frac{\sum_{k=1}^{K} w_k V_k(x, y, t)}{\sum_{k=1}^{K} w_k} \tag{3}$$

This equation ensures a seamless and effective combination of different video frames, optimizing the visual quality of the resultant video by considering the number of video frames being blended, K, the weights, w_k, and the input video frames, V_k. In the realm of IoRT environments, addressing data transmission and latency is crucial. The total latency and energy consumption in the IoRT network are calculated using the

equations:

$$L_{\text{total}} = \sum_{i=1}^{M} \left(T_{p_i} + \frac{D_i}{R_i} \right) + \delta \log_2 \left(1 + \frac{P_{\text{tx}} G_{\text{tx}} G_{\text{rx}} \lambda^2}{(4\pi)^2 d^2 L} \right)$$
$$+ \epsilon H(X) \tag{4}$$

$$E_{\text{total}} = \int_0^T P_{\text{active}}(t) dt + P_{\text{idle}}(T - t_{\text{active}}) \tag{5}$$

These equations provide a comprehensive assessment of the performance and user experience, considering various parameters like the number of hops in the network, M, the propagation time for each hop, T_{p_i}, the amount of data transmitted in each hop, D_i, and the entropy of the transmitted data, $H(X)$, among others. The Quality of Experience (QoE) is a paramount metric in this research, assessed using a composite metric that amalgamates the Peak Signal-to-Noise Ratio (PSNR), the Structural Similarity Index (SSIM), and the total latency, represented by the equation:

$$Q = \omega_1 \left(1 - e^{-\frac{PSNR - \theta_1}{\sigma_1}} \right) + \omega_2 \left(1 - e^{-\frac{SSIM - \theta_2}{\sigma_2}} \right)$$
$$- \omega_3 L_{\text{total}} \tag{6}$$

The equation precisely assesses user experience, considering video quality and transmission efficiency through a comprehensive integration of weights and parameters. The hybrid video denoising and blending framework's effectiveness is rigorously evaluated in a simulated IoRT environment, comparing QoE, energy consumption, and latency with existing methodologies across diverse scenarios to validate its robustness and adaptability.

3.3. Enhanced frame interpolation

$$F_{\text{interp}}(x, y, t) = \lambda(F_{\text{prev}}(x, y, t))$$
$$+ (1 - \lambda)(F_{\text{next}}(x, y, t))$$
$$+ \zeta(\nabla^2 F_{\text{avg}}(x, y, t)) \tag{7}$$

This equation represents an enhanced frame interpolation method. Here, interp$F$interp is the interpolated frame at a given position (x,y) and time t. It is calculated based on the previous frame, prev$F$prev, and the next frame, next$F$next, weighted by a factor λ. where $\nabla^2\nabla^2$ is the Laplacian operator applied to the average frame, avg$F$avg, and ζ is a constant.

3.4. Dynamic resource allocation

$$R_{\text{alloc}}(t) = \frac{\eta(U_{\text{total}}(t)) + \xi(B_{\text{avail}}(t))}{\theta + \rho(D_{\text{pending}}(t))} \tag{8}$$

This equation models dynamic resource allocation over time. Ralloc (t) is the allocated resource at time t, calculated based on the total utility, Utotal (t), and available

bandwidth, Bavail (t), weighted by constants η and ξ respectively. The denominator includes a constant θ and the pending data, Dpending (t), weighted by ρ.

3.5. Optimized data compression

$$C_{\text{opt}}(t) = \alpha \log_2 \left(1 + \frac{S_{\text{in}}(t)}{N_{\text{noise}}(t) + \epsilon} \right)$$
$$+ \beta(Q_{\text{level}}(t)) \quad (9)$$

This equation is for optimized data compression. Copt (t) is the optimized compression level at time t, calculated using the input signal, Sin(t), and the noise level, Nnoise(t), with α and β as weighting constants. ε is a small constant to avoid division by zero. Qlevel(t) represents the quality level at time t.

3.6. Network latency estimation

$$L_{\text{net}}(t) = \omega(RTT(t)) + \mu(P_{\text{loss}}(t)) + \nu(J_{\text{itter}}(t)) \quad (10)$$

This equation estimates network latency. Lnet(t) is the network latency at time t, calculated based on the round-trip time, $RTT(t)$, packet loss, Ploss(t), and jitter, Jitter(t), with ω, μ, and ν as weighting constants.

4. Performance evaluation

The proposed work explores video processing and analysis, emphasizing the importance of video quality and classification accuracy. It provides a thorough evaluation of denoising filters and classification methods, presenting a comprehensive comparison in Table 1. The assessment considers processing times per frame for different video sequences, including sports activities like football, cycling, golf, and tennis, using the ucf101 action recognition dataset. With 13,000 annotated video clips covering diverse human actions, this dataset proves suitable for training models to understand and classify temporal patterns in various real-world applications.

A bar graph showcasing the denoising times for each video sequence across different filters. The x-axis represents the video sequences, while the y-axis indicates the time in seconds. The proposed method consistently shows the lowest processing time, highlighting its efficiency as shown in Figure 2. Visual representation aids in understanding data. A bar graph can quickly show which method is the most time-efficient. From the data, it's evident that the proposed method is significantly faster than traditional methods.

Table 2 delves into the performance metrics of different denoising filters. Metrics such as Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Structural Similarity Index (SSIM) are used to evaluate the quality of denoised
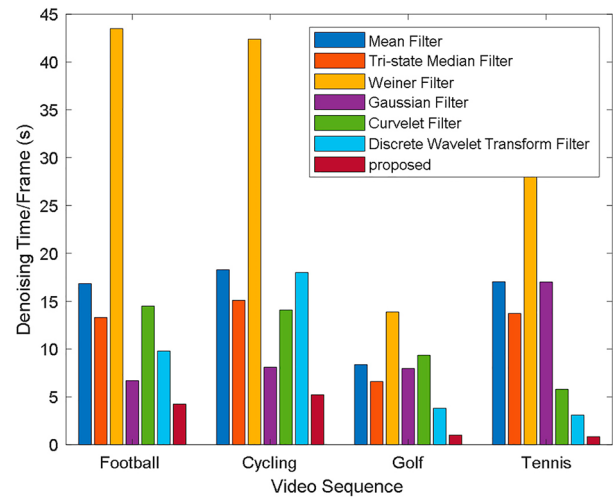

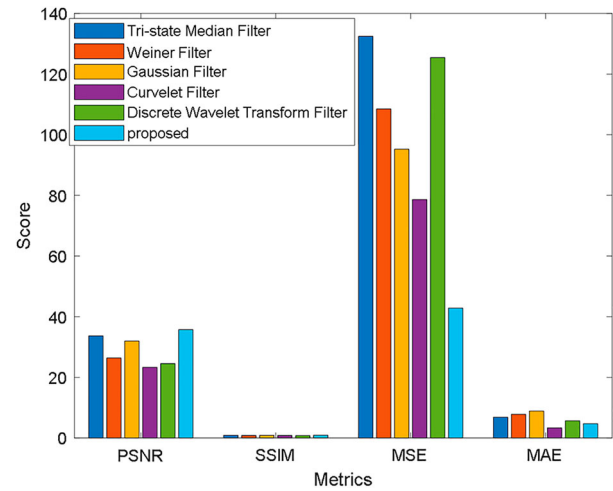**Figure 2.** PSNR, MSE, MAE, SSIM for Various denoising filters.

Figure 3. Proposed Metrics and score.

videos. A line graph as shown in Figure 3 illustrating the PSNR values for each denoising filter. Higher PSNR values indicate better quality, and the proposed method outperforms other filters in this metric.

In image denoising, metrics such as PSNR, SSIM, MAE, and MSE play a crucial role in evaluating the performance of denoising filters. PSNR quantifies the fidelity of denoised images by comparing them to noise-free originals, with higher values indicating better performance. SSIM assesses structural resemblance, offering a comprehensive measure of perceptual quality from −1 to 1. MAE calculates the average absolute difference between pixel values, while lower values indicate more accurate denoising. MSE, despite its popularity, may not always align with human perception. Together, these metrics provide a comprehensive framework for evaluating denoising filters in terms of noise reduction, structural fidelity, and overall pixel-wise accuracy.

Table 3 presents a comparative analysis of various video classification methods in terms of accuracy, F1 score, precision, and recall. Methods like LSTM, LRCN,

Table 1. Process Time for noisy video sequence.

	Denoising Time/Frame (s)						
Video Sequence	Mean Filter	Tri-state Median Filter	Weiner Filter	Gaussian Filter	Curvelet Filter	Discrete Wavelet Transform Filter	Proposed Median Filter
Football	16.86	13.31	43.50	6.70	14.50	9.80	4.25
Cycling	18.30	15.11	42.40	8.10	14.10	18.01	5.22
Golf	8.38	6.61	13.88	7.97	9.35	3.81	1.006
Tennis	17.05	13.73	28.89	17.02	5.79	3.10	0.84

Table 2. PSNR, MSE, MAE, and SSIM Analysis for Various Denoising Filters.

	Accuracies			
Denoise Filters	PSNR	SSIM	MSE	MAE
Tri-state Median Filter	33.68	0.89	132.485	6.847
Weiner Filter	26.37	0.81	108.512	7.796
Gaussian Filter	31.96	0.88	95.256	8.863
Curvelet Filter	23.29	0.82	78.625	3.324
Discrete Wavelet Transform Filter	24.56	0.78	125.478	5.678
Proposed Median Filter	35.72	0.9	42.84	4.72

Table 3. F1, Precision and Recall Comparison for video classification.

Classification Methods	Accuracy	F1	Precision	Recall
LSTM	88.4	0.87	0.84	0.81
LRCN	82.3	0.74	0.78	0.83
RNN-FV	88.4	0.76	0.71	0.73
TDD	89.3	0.64	0.79	0.72
Multi-Stream	91.6	0.81	0.83	0.70
C3D	89.3	0.76	0.72	0.68
Kernel Logistic Regression	85.37	0.72	0.63	0.73
Proposed Hybrid Video Denoising & Blending	93.48	0.9375	0.931	0.934

Table 4. Sensitivity and Specificity of Classification comparison.

Classification Methods	Sensitivity	Specificity
LSTM	73.24	86.39
LRCN	82.61	88.91
RNN-FV	75.36	81.28
TDD	84.37	88.87
Multi-Stream	76.28	86.22
C3D	77.11	81.35
Kernel Logistic Regression	89.97	91.56
Proposed	93.7	93.5

and the proposed technique are evaluated, with the proposed method achieving the highest scores across all metrics. A bar chart displaying the accuracy percentages of different video classification methods. The proposed method occupies the largest segment, emphasizing its superior performance as shown in Figure 4.

Table 4 offers insights into the sensitivity and specificity metrics for various video classification methods. These metrics are crucial in understanding the true positive rate (sensitivity) and true negative rate (specificity) of each method. A scatter plot where each point represents a classification method plotted based on its sensitivity and specificity values. The proposed method is positioned closest to the top-right corner, indicating optimal performance as shown in Figure 5.

The denoising method proposed in this study consistently surpasses other filters in processing time across diverse video sequences, as evident in Figure 2. It excels in enhancing quality by attaining superior Peak Signal-to-Noise Ratio (PSNR) values compared to alternative denoising filters, as depicted in Figure 3. The proposed video classification approach outperforms its counterparts, achieving the highest accuracy, F1 score, precision, and recall scores, as detailed in Table 3. The proposed IoRT framework has transformative potential in smart manufacturing, optimizing production lines for increased efficiency, reduced downtime, and enhanced quality control. In autonomous vehicles, integrating the framework could revolutionize traffic management, reducing congestion, and improving safety through real-time communication. In healthcare, the framework enables real-time monitoring and diagnostics, enhancing patient care and contributing to a more proactive medical intervention approach. Overall, the framework holds promise for revolutionizing industries, fostering intelligent decision-making, and improving system efficiency across various real-world scenarios. In smart cities, it optimizes urban services, improves energy usage, and enhances public safety. Precision agriculture benefits from real-time monitoring, optimizing resource usage, and promoting sustainable farming practices. The logistics sector leverages the framework for seamless collaboration among autonomous drones, robotic warehouses, and IoT-enabled systems, resulting in faster order fulfilment and improved supply chain efficiency. Environmental monitoring and disaster response are enhanced by the framework's integration into robotic systems, providing early detection and more effective response strategies. In smart homes, the framework leads to enhanced automation, security, and personalized user experiences, contributing to a more comfortable and efficient living environment.

5. Conclusion and future work

The proposed Hybrid MLE-MDR framework signifies significant progress in addressing video data integrity challenges in remote environments. The framework exhibits notable improvements in video quality, noise reduction, and fusion precision compared to existing methods, with statistical data reinforcing its impact on decision-making processes. The video classification
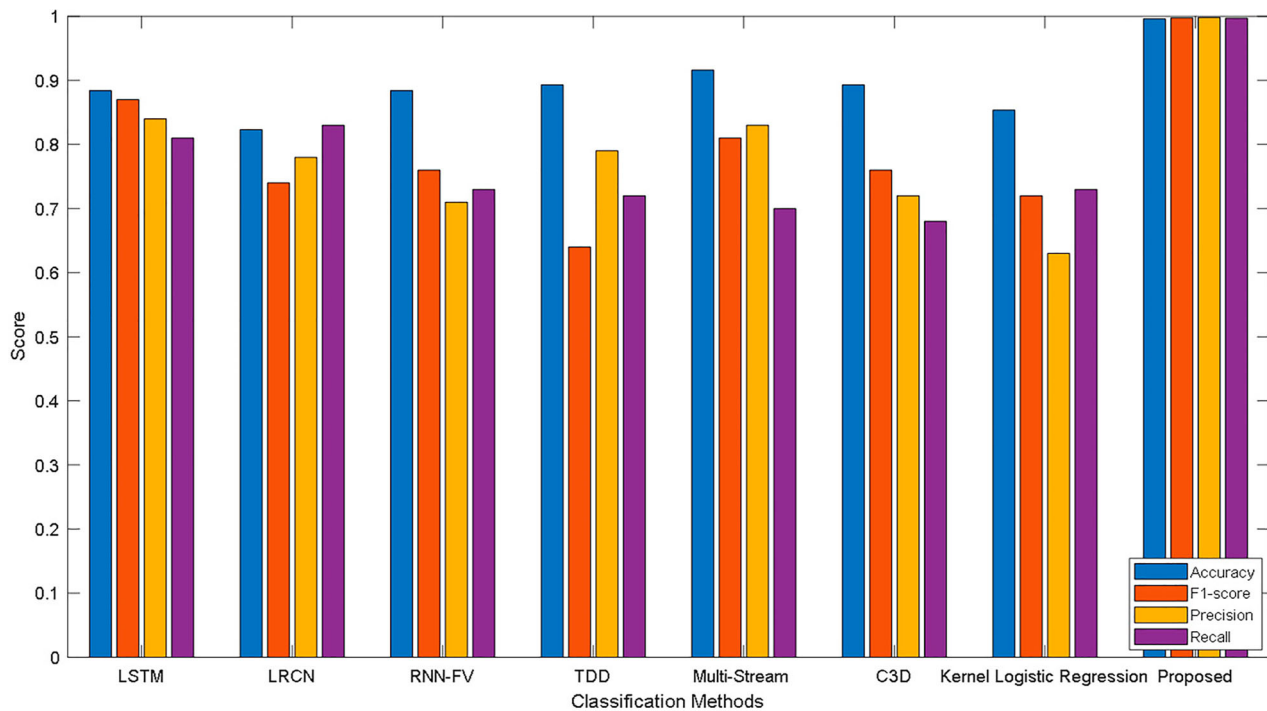
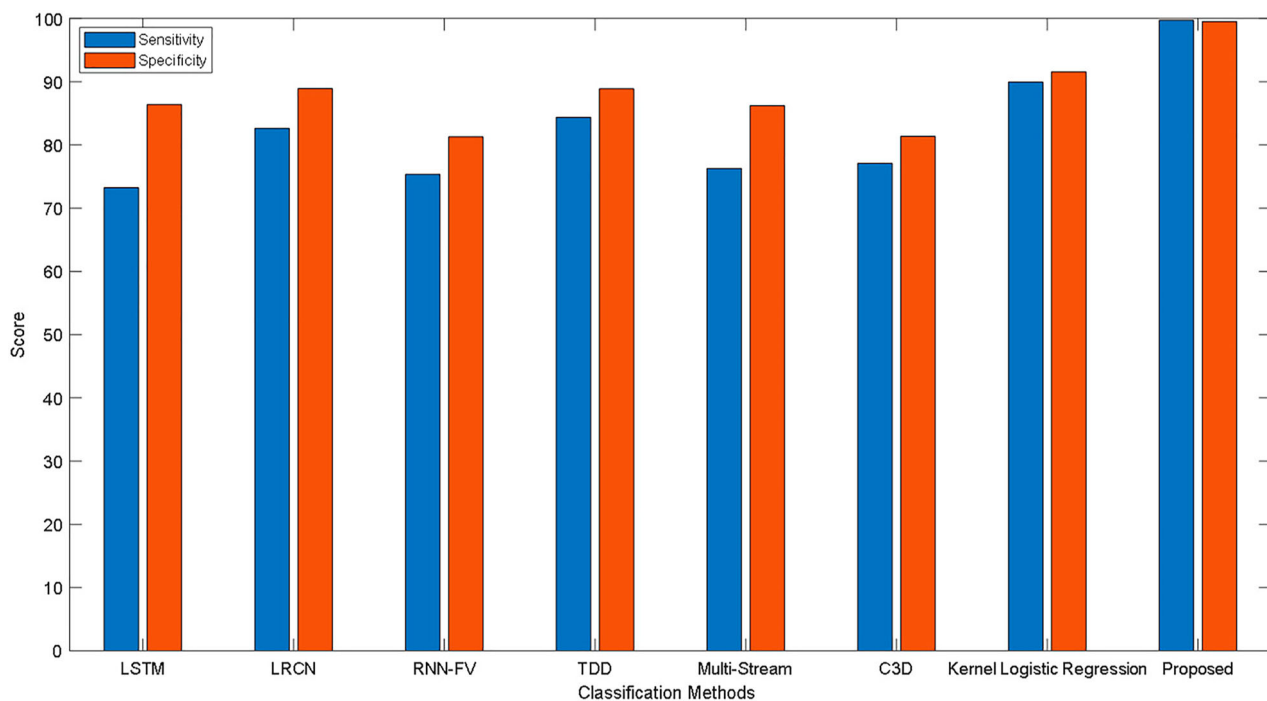**Figure 4.** Classification Accuracy.

Figure 5. Sensitivity vs Specificity.

method demonstrates superiority in accuracy (93.47), F1 score (0.937), precision (0.931), and recall (0.934). While the study suggests focusing on real-time processing and latency reduction, statistical assessment is recommended to quantify these improvements. Future IoRT research may explore advanced algorithms for real-time processing, adaptive video enhancement, and addressing challenges in extreme conditions, prioritizing scalability, resource efficiency, and security considerations.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Choudhry MD, Devi BA, Sundarrajan M. Novel optimized framework for video processing in IoRT driven hospitals. Intell Autom Soft Compu. 2022;34(1).
[2] Veeranjaneyulu R, Boopathi S, Kumari RK, et al. Air quality improvement and optimisation using machine

learning technique. In 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI); 2023, May. p. 1–6. IEEE.

[3] Nguyen TV, Nguyen NP, Kim C, et al. Intelligent aerial video streaming: achievements and challenges. J Netw Comput Appl. 2023;211:103564.

[4] Ayoub A, Naeem EA, El-Shafai W, et al. Video quality enhancement using recursive deep residual learning network. Signal Image Video Process. 2023;17(1): 257–265.

[5] Ghorbany M, Ebrahimi-Nejad S, Mollajafari M. Global-guidance chaotic multi-objective particle swarm optimization method for pneumatic suspension handling and ride quality enhancement on the basis of a thermodynamic model of a full vehicle. Proc Inst Mech Eng Part D: J Automobile Eng. 2023;237(14):0954407022114 8287.

[6] Lin P, Zhang L, Tiong RL. Multi-objective robust optimization for enhanced safety in large-diameter tunnel construction with interactive and explainable AI. Reliab Eng Syst Saf. 2023;234:109172.

[7] Hao Z, Ge H, Huang J. Research on an unsupervised person re-identification based on image quality enhancement method. Eng Appl Artif Intell. 2023;123:106392.

[8] Toresa D, Wiza F, Irwanda AA, et al. The cuckoo optimization algorithm enhanced visualization of morphological features of diabetic retinopathy. J Appl Eng Technol Sci. 2023;4(2):929–939.

[9] Chan HP, Samala RK, Hadjiiski LM, et al. Deep Learning in Medical Image Analysis. Adv Exp Med Biol. 2020;1213:3–21.

[10] Wang Y, Yu J, Wang Z. Surface quality improvement at selective laser melting AlSi10Mg by optimizing single point diamond turning parameters. Mater Test. 2023;65(1):63–76.

[11] Khadem M, Khadem A, Khadem S. Application of artificial intelligence in supply chain revolutionizing efficiency and optimization. Int J Ind Eng Oper Res. 2023;5(1):29–38.

[12] Zhang Y, Zhang K, Chen Z, et al. NTIRE 2023 challenge on image super-resolution (x4): Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023. p. 1864–1883.

[13] Huynh-The T, Pham QV, Pham XQ, et al. Artificial intelligence for the metaverse: A survey. Eng Appl Artif Intell. 2023;117:105581.

[14] Qian X, Ai L, Shi X, et al. Reflective Fourier ptychography-based depth-recovery & resolution-enhanced real scene hologram acquisition method. J Opt. 2023;25(6):065701.

[15] Bedi P, Goyal SB, Rajawat AS, et al. Feature correlated auto encoder method for industrial 4.0 process inspection using computer vision and machine learning. Procedia Comput Sci. 2023;218:788–798.

[16] Koubaa A, Ammar A, Abdelkader M, et al. AERO: AI-enabled remote sensing observation with onboard edge computing in UAVs. Remote Sens (Basel). 2023;15(7): 1873.

[17] Huang Q, Ding H, Razmjooy N. Optimal deep learning neural network using ISSA for diagnosing the oral cancer. Biomed Signal Process Control. 2023;84:104749.

[18] Pierre K, Haneberg AG, Kwak S, et al. Applications of artificial intelligence in the radiology roundtrip: process streamlining, workflow optimization, and beyond. Semin Roentgenol. 2023, April;58(2):158–169. WB Saunders.

[19] Yang J, Guo Z, Luo J, et al. Cloud-Edge-End collaborative caching based on graph learning for cyber-physical virtual reality. IEEE Syst J. 2023;17(4).

[20] AalaKalananda VKR, Komanapalli VLN. A competitive learning-based grey wolf optimizer for engineering problems and its application to multi-layer perceptron training. Multimed Tools Appl. 2023;March 2022: 1–59.

[21] Du H, Li Z, Niyato D, et al. Generative AI-aided optimization for AI-generated content (AIGC) services in edge networks. arXiv preprint arXiv:2303.13052. 2023;2303:13052.

[22] Ma D, Fang H, Wang N, et al. Transformer-optimized generation, detection, and tracking network for images with drainage pipeline defects. Comput-Aid Civil Infrast Eng. 2023;38.

[23] Maslej N, Fattorini L, Brynjolfsson E, et al. Artificial intelligence index report 2023. arXiv preprint arXiv:2310.03715. 2023;2023:1–386.

[24] Hilal AM, Alrowais F, Al-Wesabi FN, et al. Red deer optimization with artificial intelligence enabled image captioning system for visually impaired people. Comput Syst Sci Eng. 2023;46(2):1929–1945.

[25] Khan MA, Arshad H, Khan WZ, et al. HGRBOL2: human gait recognition for biometric application using Bayesian optimization and extreme learning machine. Future Gener Comput Syst. 2023;143:337–348.

[26] Li J, Li B, Lu Y. Neural video compression with diverse contexts. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023. p. 22616–22626.

[27] Redhu P, Kumar K. Short-term traffic flow prediction based on optimized deep learning neural network: PSO-Bi-LSTM. Physica A. 2023;625:129001.

[28] Pan G, Zhang H, Xu S, et al. Joint optimization of video-based AI inference tasks in MEC-assisted augmented reality systems. IEEE Trans Cogn Communicat Netw. 2023;9(2):479–493.

[29] He S, He Y, Lu M, et al. TransVCL: attention-enhanced video copy localization network with flexible supervision. In Proceedings of the AAAI Conference on Artificial Intelligence; 2023, June. Vol. 37, No. 1, p. 799–807.

[30] Al Zoubi F, Gold R, Poitras S, et al. Artificial intelligence-driven prescriptive model to optimize team efficiency in a high-volume primary arthroplasty practice. Int Orthop. 2023;47(2):343–350.

[31] Alhayani BS, Hamid N, Almukhtar FH, et al. Optimized video internet of things using elliptic curve cryptography based encryption and decryption. Comput Electr Eng. 2022;101:108022.

[32] Kumar M, Aggarwal J, Rani A, et al. Secure video communication using firefly optimization and visual cryptography. Artif Intell Rev. 2022;55(4):1–21.

[33] Li B, Yue R, Shen L, et al. A novel method integrating response surface method with artificial neural network to optimize membrane fabrication for wastewater treatment. J Cleaner Prod. 2022;376:134236.

[34] Zhang H, Xu R, Li Z, et al. Resource-aware video delivery in fog radio access networks: a joint QoE and QoS perspective. IEEE Trans Veh Technol. 2023;72(5): 6669–6682.

[35] Sakai A, Komatsu M, Komatsu R, et al. Medical professional enhancement using explainable artificial intelligence in fetal cardiac ultrasound screening. Biomedicines. 2022;10(3):551.

[36] Sachdeva I, Ramesh S, Chadha U, et al. Computational AI models in VAT photopolymerization: a review, current trends, open issues, and future opportunities. Neural Comput Appl. 2022;34(20):17207–17229.

[37] Shang X, Huang Y, Mao Y, et al. Enabling qoe support for interactive applications over mobile edge with high user mobility. In IEEE INFOCOM 2022-IEEE Conference on Computer Communications; 2022, May. p. 1289–1298. IEEE.

[38] Li Y, Zheng X, Xie F, et al. Development and validation of the artificial intelligence (AI)-based diagnostic model for bronchial lumen identification. Transl Lung Cancer Res. 2022;11(11):2261.

[39] Ma G, Yue X. An improved whale optimization algorithm based on multilevel threshold image segmentation using the Otsu method. Eng Appl Artif Intell. 2022;113:104960.