Journal of Internet of Things and Information Technology
Volume 8, Issue 2, May- August, 2025
ISSN: 3107-4502 (Online)

Ai-Driven Video Analytics in IoT for Healthcare and Agricultural Monitoring Applications

Ajay Malhotra
Professor
Department of Computer Science Engineering
IET Bhaddal, Technical Campus, Ropar, Punjab
Email id: *- ajaymalohtra352@gmail.com*

ABSTRACT

This paper explores the integration of AI-driven video analytics with Internet of Things (IoT) platforms for two high-impact domains: healthcare (patient monitoring, fall detection, ward safety) and agriculture (crop / livestock monitoring, pest/disease detection). We review recent advances in edge AI, lightweight deep learning models, and hybrid edge–cloud architectures; describe typical system architectures and processing pipelines; present representative use-cases and short case summaries; and discuss technical, ethical, and deployment challenges. The paper closes with practical recommendations and future research directions to improve robustness, privacy, and scalability.

KEYWORDS*: AI-driven video analytics, Internet of Things (IoT), edge AI, healthcare monitoring, precision agriculture, privacy-preserving inference, federated learning.*

INTRODUCTION

Rapid advancements in IoT and computer vision are enabling real-time, automated analytics of video streams for applications ranging from patient monitoring to crop-health assessment. Embedding AI inference within IoT devices—commonly referred to as edge AI—empowers deployments in sensitive environments (like hospitals) or remote locations (e.g., farms) by reducing latency, enhancing privacy, and lowering bandwidth load.

In healthcare, continuous video analytics plays a vital role in detecting falls, abnormal behaviors, and vital-sign proxies, thereby enabling early intervention and optimizing care workflows. In agriculture, video AI is used for detecting crop diseases, pests, water stress, and livestock anomalies, enabling precision agriculture and improving yield and animal welfare.

This paper integrates recent domain literature and deployment reports to outline a comprehensive system architecture, describe typical processing pipelines, illustrate real-world use cases, and identify key technical and ethical challenges. Finally, it proposes future directions to foster sustainable and trustworthy adoption at scale.

LITERATURE REVIEW

Early IoT monitoring systems relied heavily on point sensors (e.g., temperature or motion detectors) and centralized cloud analytics, leading to high latency and privacy concerns. Recent developments in model optimization (via pruning, quantization, and knowledge distillation) and efficient model backbones (e.g., MobileNet, YOLO-Lite variants) have facilitated real-time, on-device video inference.

Edge–cloud hybrid models further enable scalable architectures: preliminary or coarse inference happens at the edge, triggering cloud-based processing for deep verification or aggregation. Surveys and reviews emphasize the rollout of such systems in healthcare and agriculture, underscoring their growing relevance.

Federated learning has emerged as a privacy-preserving technique, enabling AI models to improve across distributed devices—the model aggregates updates without transferring sensitive image data. This approach is gaining traction, especially in systems where data privacy or connectivity is limited.

SYSTEM ARCHITECTURE AND PROCESSING PIPELINE

The system architecture and processing pipeline represent the backbone of AI-driven video analytics in IoT for healthcare and agricultural monitoring applications. Both components ensure that visual data flows seamlessly from collection to actionable insights, with attention to real-time constraints, privacy, and scalability.

System Architecture

A typical AI–IoT video analytics architecture can be divided into five interconnected layers:

a) Sensing Layer

This is the first interface between the physical environment and the digital system. It comprises:

- Video Sources: Fixed surveillance cameras in hospitals or fields, PTZ (Pan-Tilt-Zoom) cameras for wide coverage, or UAV-mounted cameras in agricultural lands.
- Ancillary Sensors: Temperature, humidity, soil moisture, and wearable sensors (e.g., patient heart rate, cattle movement tags) complement video streams with environmental or physiological data.
- Data Characteristics: Continuous, high-volume video streams requiring compression or intelligent frame sampling to avoid overwhelming subsequent layers.

b) Edge Processing Layer

The edge layer performs on-device AI inference using resource-constrained hardware such as Raspberry Pi with Google Coral TPU, NVIDIA Jetson Nano, or similar low-power accelerators. Key functions include:

- Pre-processing: Frame resizing, background subtraction, noise reduction, and ROI (Region of Interest) detection to reduce computational load.
- On-device Analytics: Deployment of lightweight deep learning models such as Tiny-YOLO or MobileNet-SSD for real-time object detection or event recognition.
- Advantages: Low latency, reduced bandwidth usage, and enhanced privacy since raw videos remain local unless necessary to transmit.

c) Connectivity Layer

This layer ensures reliable communication between the edge devices, cloud servers, and user applications:

- Communication Standards: LoRaWAN or NB-IoT for low-bandwidth remote farms; Wi-Fi or 4G/5G for hospitals with higher bandwidth demands.
- Challenges: Network interruptions and variable bandwidth require adaptive streaming strategies, such as transmitting only critical event snippets rather than entire video streams.

d) Cloud/Analytics Layer

The cloud layer handles centralized analytics, long-term storage, and AI model management.

- Functions:
    - Aggregates data from multiple edge devices.
    - Performs advanced analytics using larger deep learning models or ensemble methods.
    - Retrains models periodically using federated learning or transfer learning approaches.
- Scalability: Supports hundreds of distributed edge devices while maintaining a unified dashboard for monitoring.

e) Application Layer

This is the end-user interaction interface, providing:

- Alerts: Real-time notifications (e.g., SMS, app notifications) for detected events like patient falls or crop diseases.
- Visualization Dashboards: Temporal trends, anomaly heatmaps, and video clips for decision-making.
- Decision-Support Tools: Automated irrigation control in agriculture or nurse allocation in healthcare wards.

Processing Pipeline

The processing pipeline is a step-by-step workflow converting raw video into meaningful alerts and insights. Each stage applies specific techniques to handle scale, accuracy, and privacy constraints.

Step 1: Input Video Acquisition

Continuous video streams originate from multiple cameras or UAVs. Multi-camera synchronization and temporal alignment are often required to ensure event consistency across sources.

Step 2: Frame Sampling

To reduce redundancy, only selected frames or video segments are analyzed. Adaptive sampling techniques dynamically adjust frame rates based on motion detection or environmental triggers (e.g., low activity periods → fewer frames analyzed).

Step 3: Pre-Processing

Pre-processing enhances raw frames before feeding them to AI models:

- ROI extraction focuses on regions with maximum information (e.g., patient bed area, crop canopy).
- Compression-aware transforms balance quality and bandwidth by using formats like H.265 with minimal visual degradation.
- Lighting normalization compensates for illumination variations in outdoor farms or hospital corridors.

Step 4: Edge Inference

Lightweight deep learning models deployed on the edge classify or detect objects/events in real time. For example:

- Healthcare: Fall detection using pose estimation models.
- Agriculture: Leaf disease classification using MobileNet-based CNNs. Only key metadata (bounding boxes, class labels) or compressed video clips are transmitted further.

Step 5: Temporal Aggregation

Events are tracked over time to avoid false positives. For instance, livestock behavior anomalies require multiple consecutive frames indicating lameness rather than a single noisy detection.

Step 6: Trigger Logic

Rules or AI-based decision modules determine whether to alert users or request higher-level cloud analysis. Example:

- Healthcare: Trigger alerts only if a fall is detected with >90% confidence.
- Agriculture: Trigger irrigation if drought stress persists across multiple time windows.

Step 7: Cloud Processing

The cloud performs:

- Deeper inference using computationally heavy models (e.g., 3D-CNNs, Vision Transformers).
- Long-term analytics to identify temporal patterns (e.g., seasonal crop health trends).

- Model updates via federated learning to keep edge models current.

Step 8: Action and Visualization

Finally, actionable insights reach end-users through:

- Mobile/Web Dashboards: Summarizing detections and trends.
- Automated Actuation: Irrigation systems or patient alarm triggers.
- Decision-Support Recommendations: Suggestions for crop treatment or patient check-ups.

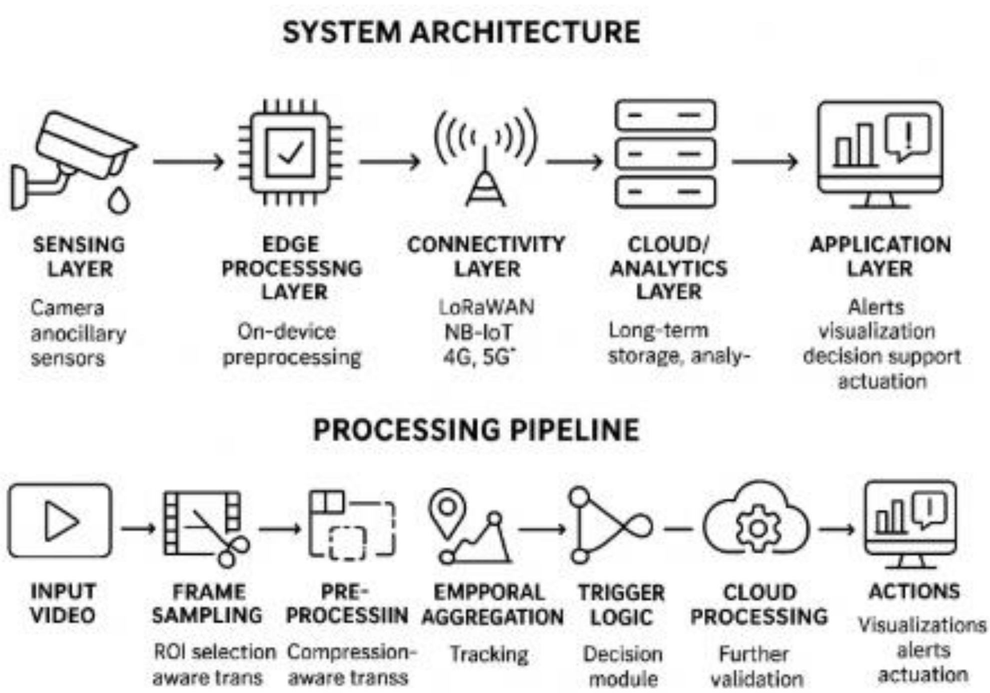

Figure: 1 System Architecture and Processing Pipeline

CASE STUDIES & APPLICATION SCENARIOS

4.1 Healthcare: Remote Patient and Ward Monitoring

- **Use Case:** Edge devices continuously monitor patient rooms or wards for fall risk and abnormal activity. Upon detecting a fall or unusual behavior, encrypted video snippets or bounding box metadata are forwarded to healthcare staff.
- **Advantages:** Real-time alerts enhance patient safety; edge-first design reduces exposure of sensitive data and minimizes network bandwidth usage.
- **Evidence:** Hospitals are piloting such systems to improve staff allocation and patient oversight, citing reduced response latency and improved data privacy.

Journal of Internet of Things and Information Technology
Volume 8, Issue 2, May- August, 2025
ISSN: 3107-4502 (Online)

Agriculture: Crop and Livestock Monitoring

- **Use Case A – Crop Health:** Fixed or aerial cameras detect early disease or pest outbreaks by recognizing leaf lesions, discoloration, or abnormal canopy patterns. On-device inferences guide targeted treatment.
- **Use Case B – Dairy Livestock Welfare:** Video-based gait and posture analytics detect lameness or illness in cattle, supplemented with environmental sensor data. Alerts to farmers enable early intervention, improving animal welfare and productivity.
- **Evidence:** Academic and industry deployments—especially in India—show promising results in detecting disease or behavioral anomalies, reducing manual labor and improving response time to issues.

CHALLENGES

1. **Data Quality & Variability:** Changing lighting, occlusion, weather, and heterogeneous subject appearances (e.g., different patient groups or crop types) require robust domain adaptation.
2. **Connectivity & Energy Constraints:** Remote sites or resource-limited clinics demand ultra-efficient edge models with intermittent connectivity and low power usage.
3. **Privacy, Ethics & Regulation:** Especially in healthcare, video often captures identifying information. Complying with privacy regulations and ensuring fairness across demographic subgroups is critical.
4. **Model Drift & Maintenance:** Seasonal changes in crops or demographic shifts in patient populations can degrade model performance; continuous retraining and domain adaptation are necessary.
5. **Interpretability & Adoption:** End users—clinicians and farmers—need transparent AI decisions with confidence estimates, explanations, and clear thresholds to trust and adopt these systems.

FUTURE DIRECTIONS

- **Federated and Collaborative Learning:** Distribute model updates across hospitals or farms without centralizing sensitive video data, enhancing generalization while preserving privacy.

- **Multimodal Fusion:** Combine video analytics with thermal imaging, acoustic sensors, vibration detectors, and wearable data to improve robustness and reduce false positives.
- **Hardware–Software Co-Design:** Leverage TinyML, NPUs, and low-power accelerators tailored for video inference to reduce latency and energy draw.
- **Explainability and Ethical Toolkits:** Incorporate interpretability methods (e.g., heatmaps, decision summaries) and bias detection tools to build trust and align with emerging regulations.
- **Edge–Cloud Orchestration:** Develop adaptive pipelines that dynamically shift computation between edge and cloud based on network conditions, energy constraints, and urgency.

CONCLUSION

Integrating AI-driven video analytics with IoT infrastructure unlocks powerful capabilities for healthcare and agricultural monitoring, offering timely insights with reduced bandwidth demands and enhanced data privacy. Although pilot deployments demonstrate substantial benefits, scalable implementation hinges on overcoming data variability, resource constraints, privacy concerns, and maintaining user trust. Future research, combining federated learning, multimodal sensing, hardware acceleration, and interpretability, will be essential to realize scalable, equitable, and resilient monitoring solutions.

REFERENCES

1. “The IoT and AI in Agriculture: The Time Is Now – Systematic review.” *PMC*.
2. “A Survey on Video Big Data Analytics: Architecture, Technologies …” *Applied Sciences (MDPI)*.
3. “How Edge AI Can Maximize Resources and Boost Productivity.” Intel Whitepaper.
4. “Fairness and Ethics in AI-Driven Healthcare: A Survey.” *PMC* (2024).
5. “AI-IoT Based Smart Agriculture — Plant Disease Detection.” *Scientific Reports* (2025).
6. “IIIT-A Develops Smart Dairy Monitoring System.” *The Times of India* (2025).