



Video for Computer Vision in Defence Applications

A Strategic Whitepaper by Astute Systems

3 January 2026



The Critical Role of Video for Computer Vision in Modern Defence Applications



Real time video in defence for C2 and C4ISR

Executive Summary

The modern battlefield is increasingly defined by data, and high-fidelity video streams are at the core of this transformation. As defence systems evolve towards greater autonomy, enhanced situational awareness, and networked operations, the integration of video for computer vision (CV) applications becomes paramount. This whitepaper, authored by Astute Systems, a leading UK-based technology provider in defence solutions, delves into the technical requirements and suitability of various video standards for defence. We critically examine the trade-offs between network-centric approaches like RTP over Ethernet, as leveraged by DEF-STAN 00-082 (Generic Vehicle Architecture - GVA), and the low-latency, high-reliability characteristics of SDI for mission- and safety-critical systems. By providing detailed specifications and operational insights, Astute Systems aims to guide technical decision-makers and engineers in navigating the complexities of video integration, ensuring optimal performance, interoperability, and security for the next generation of defence platforms.

Introduction & Background

The landscape of modern defence operations is undergoing a profound shift, driven by advancements in sensor technology, artificial intelligence (AI), and networked command and control (C2) systems. At the heart of this evolution lies video—not merely as a display medium, but as a fundamental data source for

sophisticated computer vision (CV) applications. From autonomous navigation and target recognition to enhanced situational awareness and predictive maintenance, the ability to capture, transmit, process, and interpret video data in real-time is becoming a decisive factor in operational effectiveness.

Historically, video in defence applications primarily served as a direct visual aid for human operators, often relying on analog signals or basic digital point-to-point connections. However, the escalating demand for automated threat detection, precise weapon guidance, and complex robotic operations necessitates a paradigm shift. Today, video feeds are increasingly processed by algorithms that can identify patterns, track objects, and even make autonomous decisions, often faster and with greater consistency than human operators. This transition places stringent demands on video standards, requiring not only high resolution and frame rates but also exceptionally low latency, robust signal integrity, and seamless integration into overarching system architectures.

The proliferation of diverse sensors—electro-optical (EO), infrared (IR), thermal, and multi-spectral—generates an immense volume of video data. Effectively harnessing this data for computer vision requires a deep understanding of underlying video transmission technologies, their inherent limitations, and their suitability for specific operational contexts. This paper will dissect these critical technical considerations, offering a comprehensive analysis to inform strategic decisions in defence procurement and system design.

Current State Analysis

The defence industry is actively pursuing architectures that promote modularity, scalability, and interoperability. A prime example of this is the NATO Generic Vehicle Architecture (GVA), codified in DEF-STAN 00-082, which provides a framework for standardizing the electrical and electronic systems within military vehicles. The GVA mandates an Ethernet-based network backbone for data exchange, including video, aiming to reduce integration costs, simplify upgrades, and enhance crew efficiency.

As the Luminact GVA White Paper (2017) highlights, "The Generic Vehicle Architecture (GVA) standard, as detailed in DEF-STAN 00-082, is a pivotal development aimed at standardizing electrical and electronic architectures within military vehicles, facilitating the integration of disparate systems and sensors, including video feeds." This network-centric approach allows for flexible routing of video streams from any sensor to any display or processing unit within the vehicle, and potentially to external networks for broader situational awareness or remote command.

The current state sees a diverse array of video sources, from high-definition EO/IR cameras on unmanned aerial vehicles (UAVs) to panoramic driver vision enhancer (DVE) systems on armoured vehicles. These systems produce video data at varying resolutions, frame rates, and compression levels. The challenge lies in harmonizing these disparate feeds into a cohesive system that can reliably deliver data for real-time computer vision applications, often under extreme environmental conditions and stringent security requirements.

The shift towards AI and machine learning (ML) in defence further amplifies the need for high-quality video. CV algorithms, whether for object classification, anomaly detection, or predictive analytics, perform optimally with clear, uncorrupted, and synchronized video inputs. This necessitates a careful evaluation of video transmission protocols, considering factors like bandwidth efficiency, latency performance, and resilience to network degradation.

Key Challenges & Opportunities

The integration of video for computer vision in defence presents a complex array of challenges alongside significant opportunities for enhancing operational capabilities.

Challenges:

- **Latency and Real-time Processing:** Many defence CV applications, such as weapon targeting, autonomous navigation, and close-quarters manoeuvring, demand near-zero latency. Encoding, transmission over networks, and decoding processes inherently introduce delays, which can be unacceptable for critical real-time decision-making.
- **Bandwidth Management:** High-resolution, high-framerate video generates enormous amounts of data. Transmitting multiple such streams across a networked architecture, especially over wireless links or congested vehicle networks, can quickly saturate available bandwidth, leading to dropped frames, reduced quality, or increased latency.
- **Interoperability and Standardization:** Defence platforms often comprise systems from various manufacturers, leading to proprietary interfaces and protocols. While standards like GVA aim to address this, ensuring seamless interoperability of video streams for CV across different generations of equipment and international partners remains a significant hurdle.
- **Environmental Robustness and Reliability:** Defence systems operate in harsh environments, characterized by extreme temperatures, shock, vibration, and electromagnetic interference (EMI). Video transmission infrastructure must be ruggedized to ensure continuous operation and data integrity under these conditions.
- **Security:** Video feeds, especially those processed by CV algorithms, can contain sensitive information. Protecting these streams from interception, spoofing, or tampering is critical. Cybersecurity measures must be integrated at every layer, from sensor to display and processing unit.
- **Data Fusion and Synchronization:** CV applications often rely on fusing video data with inputs from other sensors (e.g., LiDAR, radar, GPS). Accurate time synchronization across all these data streams is essential for precise object localization and situational awareness, adding complexity to system design.

Ross Newman, Director at Astute Systems notes that "The integrity and reliability of video data are non-negotiable for defence computer vision applications. A compromised or delayed video feed can lead to catastrophic failures, making robust transmission and processing architectures paramount."

Opportunities:

- **Enhanced Situational Awareness:** CV algorithms can process vast amounts of video data much faster than human operators, identifying threats, tracking multiple targets, and providing a comprehensive, fused operational picture, thereby reducing cognitive load and improving decision-making speed.
- **Autonomous Operations:** High-fidelity video combined with advanced CV is fundamental for enabling truly autonomous systems, from self-driving vehicles and robotic reconnaissance units to automated target engagement and drone swarms.
- **Predictive Maintenance:** CV can monitor the condition of equipment, identifying wear and tear or anomalies in real-time, enabling predictive maintenance schedules that reduce downtime and extend asset life.
- **Improved Training and Simulation:** Realistic video feeds from operational environments, processed by CV, can be used to generate highly accurate training simulations, preparing personnel for complex scenarios.
- **Reduced Manpower and Risk:** Automation through CV can reduce the need for human presence in hazardous environments, lowering operational costs and minimizing risk to personnel.

Detailed Analysis

1. Video Standards and Their Suitability for Defence Applications

The choice of video transmission standard is a critical architectural decision, impacting performance, cost, and system longevity. Each standard offers a unique set of characteristics regarding resolution, framerate, cable requirements, connector types, and, crucially, latency. The following table provides a comparative overview of key video standards relevant to defence, highlighting their strengths and typical use cases.

| Feature | SDI (e.g., 3G-SDI, 12G-SDI) | RTP (over Ethernet) | HDMI (Consumer/Prosumer) | DisplayPort (DP) | CoaXPress (CXP) |
|--------------|--|--|-------------------------------------|------------------------------------|--|
| Resolution | Up to 8K UHD (12G-SDI) | Varies (codec dependent, e.g., 4K, 8K) | Up to 8K UHD | Up to 16K UHD | Up to 8K UHD |
| Framerate | Up to 60fps (4K), 120fps (HD) | Varies (codec dependent, e.g., 60fps, 120fps) | Up to 120Hz (4K), 240Hz (HD) | Up to 240Hz (4K), 360Hz (HD) | Up to 60fps (4K) |
| Cable Length | ~100m (coaxial, depending on resolution/cable quality) | ~100m (Cat5e/6), kilometers (fiber optic) | ~15m (copper), >100m (active/fiber) | ~3m (copper), >100m (active/fiber) | ~100m (coaxial, up to 12.5 Gbps) |
| Connectors | BNC (75 Ohm) | RJ45, Fiber Optic (LC, ST, SC) | Type A, Type D (Micro) | DP, USB-C | BNC, DIN 1.0/2.3 |
| Latency | Very Low (<1ms typical, uncompressed) | Variable (ms to hundreds of ms, codec/network dependent) | Low (typically <10ms) | Low (typically <10ms) | Very Low (<1ms typical, uncompressed) |
| Key Use Case | Mission/Safety-critical, direct control, broadcast | Networked systems, GVA, distributed C2, ISR | Displays, non-critical AV systems | High-res displays, multi-monitor | High-speed machine vision, industrial automation |

2. RTP Video and DEF-STAN 00-082 for Ethernet-based Streaming

Real-time Transport Protocol (RTP) is a widely adopted standard for streaming audio and video over IP networks. Its application within military platforms, particularly under the guidance of DEF-STAN 00-082's Generic Vehicle Architecture (GVA), represents a significant shift towards network-centric operations.

Benefits of RTP Video in GVA Environments:

- **Network-Centricity and Interoperability:** RTP over Ethernet aligns perfectly with the GVA's core principle of a unified digital backbone. It enables any video source to connect to any processing unit or display via a standardized IP network, facilitating modularity and easing integration. The Luminact GVA White Paper (2017) emphasizes that "The GVA architecture defines a common digital backbone for military vehicles, facilitating the integration of disparate systems and sensors, including video feeds." This allows for flexible routing and sharing of video streams across multiple vehicle systems and even external networks.
- **Scalability and Distribution:** IP-based video allows for easy distribution to multiple endpoints simultaneously. A single video feed can be multicast to several displays, recording devices, or computer vision processors without requiring dedicated point-to-point hardware for each destination.

- **Flexibility in Codec Selection:** RTP can encapsulate various video codecs (e.g., H.264, H.265, JPEG2000), allowing system designers to balance image quality, bandwidth consumption, and processing complexity. H.265 (HEVC), for instance, offers significantly better compression efficiency than H.264, crucial for transmitting high-resolution video over constrained military networks.
- **Data Integration and Fusion:** Being on an IP network, video streams can be seamlessly combined and synchronized with other sensor data (e.g., LiDAR point clouds, radar tracks, GPS coordinates) for advanced computer vision applications and comprehensive situational awareness. This data fusion is critical for complex autonomous behaviours.
- **Reduced Cabling Complexity:** A single Ethernet cable can carry multiple video streams along with other data, reducing the overall cable weight and complexity within a vehicle compared to multiple dedicated coaxial or fibre optic runs.

Disadvantages of RTP Video for Computer Vision:

- **Latency:** This is the most significant drawback for critical CV applications. Latency in RTP video stems from several sources:
- **Encoding/Decoding:** Compression codecs introduce processing delays. Even highly optimized codecs can add tens to hundreds of milliseconds.
- **Packetization:** Video frames must be broken into IP packets, adding a small delay.
- **Network Transit:** Packet travel time across the network, especially with multiple hops or wireless links.
- **Jitter Buffering:** To compensate for variable network delays (jitter), receiving devices buffer packets, adding further latency.
- **Display Processing:** The display device itself can introduce minor delays.

For applications requiring immediate feedback, such as weapon aiming or direct vehicle control, even 100-200ms of latency can be unacceptable.

- **Bandwidth Consumption:** While codecs help, uncompressed or lightly compressed high-resolution, high-framerate video still requires substantial bandwidth. Congestion on the Ethernet network can lead to packet loss, affecting video quality or causing dropped frames, which can severely impact CV algorithm performance.
- **Network Complexity:** Implementing robust RTP video streaming requires careful network design, including Quality of Service (QoS) mechanisms to prioritize video traffic, sophisticated routing, and error recovery protocols. This complexity increases system vulnerability and management overhead.
- **Security Vulnerabilities:** As IP-based traffic, RTP streams are susceptible to standard network attacks (e.g., denial-of-service, eavesdropping, man-in-the-middle). Robust encryption and authentication protocols are essential, adding further processing overhead and potential latency.

According to Ross Newman, Director at Astute Systems, "While RTP over Ethernet offers unparalleled flexibility and aligns with modern network architectures like GVA, it introduces inherent latency that must be meticulously managed. For computer vision applications demanding sub-10ms response times, the trade-off between network flexibility and real-time performance becomes a critical design constraint."

3. SDI Video for Safety-Critical and Low-Latency Applications

Serial Digital Interface (SDI) is a broadcast industry standard that transmits uncompressed, unencrypted digital video over coaxial cable. While not inherently network-centric like RTP, its fundamental characteristics make it exceptionally well-suited for specific, safety-critical, and low-latency defence applications.

Benefits of SDI Video:

- **Extremely Low Latency:** SDI transmits video data serially and uncompressed, minimizing processing overheads. This results in glass-to-glass latency often measured in microseconds, typically less than 1ms. This near-instantaneous feedback is crucial where any perceptible delay could have severe consequences. The Luminact GVA White Paper (2017) acknowledges this, stating, "SDI's minimal processing overhead ensures negligible latency, a critical factor for direct-drive applications and weapon systems where even milliseconds can impact operational effectiveness."
- **High Reliability and Robustness:** Coaxial cables are inherently robust and less susceptible to electromagnetic interference (EMI) over short-to-medium distances compared to unshielded twisted pair (UTP) Ethernet cables. The point-to-point nature of SDI links simplifies troubleshooting and reduces potential failure points associated with complex network topologies.
- **Guaranteed Bandwidth and Quality:** Since SDI transmits uncompressed video, there are no concerns about compression artifacts, packet loss, or variable bandwidth impacting image quality. The video stream is pristine from source to destination.
- **Simplicity:** SDI connections are straightforward point-to-point links, requiring minimal configuration or network management compared to IP-based systems.
- **Proven Technology:** SDI has a long track record of reliability in demanding broadcast and industrial environments, providing a mature and stable technology for mission-critical defence applications. Modern SDI standards like 12G-SDI support 4K UHD resolution at 60fps, providing ample fidelity for advanced CV.

Use Cases for SDI Video in Defence:

- **Driving Systems (Driver Vision Enhancers - DVEs):** For armoured vehicles operating in degraded visibility or with closed hatches, DVE systems provide the driver with a panoramic or direct view of the surroundings. Any latency in these systems could lead to delayed reactions, resulting in collisions or inability to navigate complex terrain. SDI ensures that the visual feedback is immediate, enabling precise vehicle control.
- **Weapon Systems and Target Acquisition:** In weapon platforms, the video feed from a targeting sensor (e.g., EO/IR camera) to the fire control system and operator display must be instantaneous. Latency here can mean missing a fleeting target, inaccurate aiming, or even friendly fire. SDI's low latency ensures that the operator sees the target exactly as the weapon system does, enabling rapid and accurate engagement. This is critical for systems like remote weapon stations (RWS) and guided missile platforms.
- **Pilotage and Flight Control:** For manned aircraft, especially helicopters or UAVs operating in complex airspaces or performing precision landings, direct visual feedback without delay is paramount. SDI can provide the necessary low-latency video for critical flight control instruments and pilot displays, particularly from external cameras.
- **Safety-Critical Monitoring:** Any application where human life or high-value assets depend on immediate visual feedback, such as monitoring critical machinery, hazardous material handling, or explosive ordnance disposal (EOD) robotics, benefits immensely from SDI's reliability and low latency.

As Ross Newman, Director at Astute Systems observes, "While the future of defence undoubtedly leans towards networked IP systems, neglecting the unparalleled low latency and reliability of SDI for safety-critical applications would be a grave error. A hybrid architecture, intelligently leveraging both SDI and RTP, offers the optimal balance of network flexibility and operational assurance."

4. Computer Vision Integration and Processing

The efficacy of computer vision in defence applications relies not just on the video transmission standard but equally on the processing architecture.

- **Edge Computing:** With the proliferation of sensors and the need for immediate action, there's a growing trend towards processing video data at the "edge"—i.e., directly on the sensor platform or within the vehicle. This reduces bandwidth requirements for data transmission to a central server, minimizes latency, and allows for faster decision-making. Edge AI processors (e.g., NVIDIA Jetson series, Intel Movidius) are specifically designed for this purpose, performing tasks like object detection, tracking, and classification directly where the data is captured.
- **Centralized Processing:** For more complex analyses, fusion of data from multiple vehicles or platforms, and long-term data storage/retrieval, centralized processing units (either on-board larger platforms or in ground stations) are still essential. These systems can leverage more powerful GPUs and cloud-like infrastructure to run sophisticated AI/ML models, provide comprehensive situational awareness, and support strategic decision-making.
- **Data Fusion:** True situational awareness often requires fusing video data with other sensor inputs. For CV algorithms, this means correlating visual information with radar tracks, LiDAR point clouds, acoustic signatures, and geospatial data. Precise time synchronization of all these data streams is crucial to avoid misinterpretations and ensure accurate object localization and tracking.
- **AI/ML Model Development:** The performance of CV applications is directly tied to the quality and robustness of the underlying AI/ML models. This requires extensive training data, often annotated video feeds, and continuous refinement to adapt to new threats, environments, and operational conditions. The ability to collect, store, and process large volumes of high-quality video data is therefore fundamental.

Recommendations

Based on this comprehensive analysis, Astute Systems offers the following recommendations for integrating video for computer vision in defence applications:

- **Adopt a Hybrid Architecture:** Do not exclusively commit to a single video standard. Implement a hybrid approach where SDI is leveraged for all safety-critical, low-latency applications (e.g., driving, weapon systems, pilotage) and RTP over Ethernet is utilized for distributed situational awareness, C2, and less latency-sensitive CV applications within the GVA framework. This ensures both operational safety and network flexibility.
- **Prioritize Open Standards and Modular Design:** Insist on systems that adhere to open standards like GVA (DEF-STAN 00-082) and utilize standardized video codecs (e.g., H.265, JPEG2000). This promotes interoperability, reduces vendor lock-in, simplifies upgrades, and future-proofs defence platforms against evolving threats and technologies.
- **Invest in Robust Cybersecurity:** All networked video streams and CV processing units must be protected by comprehensive cybersecurity measures. This includes end-to-end encryption, strong authentication protocols, intrusion detection systems, and regular vulnerability assessments to prevent unauthorized access, data manipulation, or denial-of-service attacks.
- **Optimize for Latency and Bandwidth:** For RTP-based systems, prioritize codecs and network configurations that minimize latency while efficiently managing bandwidth. Implement Quality of Service (QoS) mechanisms to guarantee bandwidth for critical video streams. Explore techniques like region-of-interest (ROI) encoding to focus bandwidth on relevant areas.
- **Develop Edge Processing Capabilities:** Invest in robust, ruggedized edge computing hardware and software capable of performing real-time computer vision tasks directly on the sensor platform or within the vehicle. This reduces reliance on central processing, minimizes network load, and accelerates decision cycles.
- **Ensure Precise Time Synchronization:** For any CV application involving data fusion from multiple sensors, implement highly accurate time synchronization protocols (e.g., PTP - Precision Time Protocol) across all video sources and other sensors. This is critical for accurate object tracking, mapping, and decision-making.

- **Rigorous Testing and Validation:** All video and CV systems must undergo extensive testing and validation under realistic operational conditions, including stress testing for latency, bandwidth degradation, environmental resilience, and cybersecurity vulnerabilities.

Conclusion

The integration of video for computer vision is not merely an enhancement but a fundamental requirement for the next generation of defence capabilities. As militaries increasingly rely on autonomous systems, enhanced situational awareness, and rapid decision-making, the choice and implementation of video transmission standards become critically important.

Astute Systems recognizes that there is no one-size-fits-all solution. While the network-centric approach of RTP over Ethernet, as championed by DEF-STAN 00-082 and GVA, offers immense flexibility and scalability for distributed information sharing, its inherent latency makes it unsuitable for all applications. Conversely, SDI continues to be the gold standard for safety-critical, low-latency applications like driving vision and weapon systems, where immediate feedback is paramount.

The optimal path forward involves a strategic, hybrid approach that intelligently leverages the strengths of both RTP and SDI, ensuring that the right video standard is applied to the right operational requirement. This, coupled with robust cybersecurity, advanced edge processing, and meticulous system design, will unlock the full potential of computer vision in defence.

Astute Systems stands at the forefront of this technological evolution, offering deep expertise in defence technology, military systems integration, and national security solutions. We are committed to partnering with defence organizations to design, develop, and deploy cutting-edge video and computer vision systems that enhance operational effectiveness, ensure personnel safety, and maintain a decisive technological advantage. By understanding the nuanced requirements and technical specifications, we enable our clients to navigate the complexities and build resilient, future-proof defence platforms.



Get In Touch

 astutesys.com

 enquiries@astutesys.com



Suite 2.08, The Precinct, Brunswick Street, Fortitude Valley, QLD
4006