

WOWZA VIDEO INTELLIGENCE FRAMEWORK

AI-Powered Analysis for Live Video Streams

Wowza Video Intelligence Framework (VIF) enables AI-powered analysis of streaming media by extracting video frames, routing them to external AI models for inference, and converting the results into real-time outputs such as metadata, overlays, event logs, and alerts. Organizations can treat video streams as a programmable source of operational data — not a passive media feed.

Turning live video into operational intelligence — across transportation, defense, public safety, and more.

PRACTICAL APPLICATIONS & USE CASES



Department of Transportation

Camera health monitoring, wrong-way driver detection, stalled vehicle detection, infrastructure security.



Defense & Aerospace

Perimeter intrusion detection, weapons detection, crowd/mass gathering analysis, range safety surveillance.



Law Enforcement

Drone feed monitoring, inter-agency surveillance sharing, corrections facility oversight, SWAT operational support.



Industrial & Infrastructure

Equipment monitoring, zone-based anomaly detection, post-incident event logging.



Live Events & Venues

Stadium and venue security monitoring, crowd density analysis, perimeter access control, incident detection during large-scale events.



Critical Facilities

Access control monitoring, restricted area detection, camera health across distributed sites.

WHY ORGANIZATIONS CHOOSE WOWZA VIDEO INTELLIGENCE FRAMEWORK

1

No Infrastructure Rebuild Required

Wowza VIF connects to existing Wowza Streaming Engine deployments. There is no need to replace your video infrastructure to add AI capabilities.

2

Model Flexibility

Bring your own pre-trained models. Wowza VIF is designed to support a wide range of computer vision models including custom object detection (RF-DETR).

3

Real-Time Operational Signals

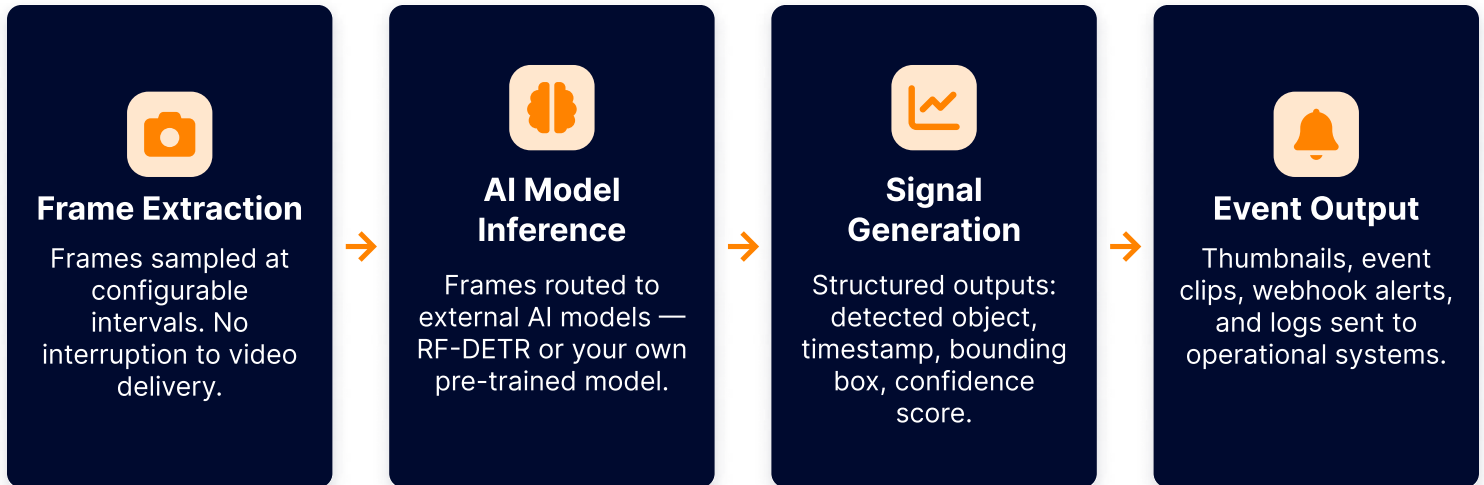
Convert AI detections into structured outputs: metadata, bounding boxes, thumbnails, event clips, confidence scores, and webhook alerts.

4

Independent Scaling

AI inference runs in a separate service so compute-intensive workloads never impact the reliability of your video delivery pipeline.

HOW WOWZA VIDEO INTELLIGENCE FRAMEWORK WORKS



WHAT WOWZA VIDEO INTELLIGENCE FRAMEWORK ENABLES

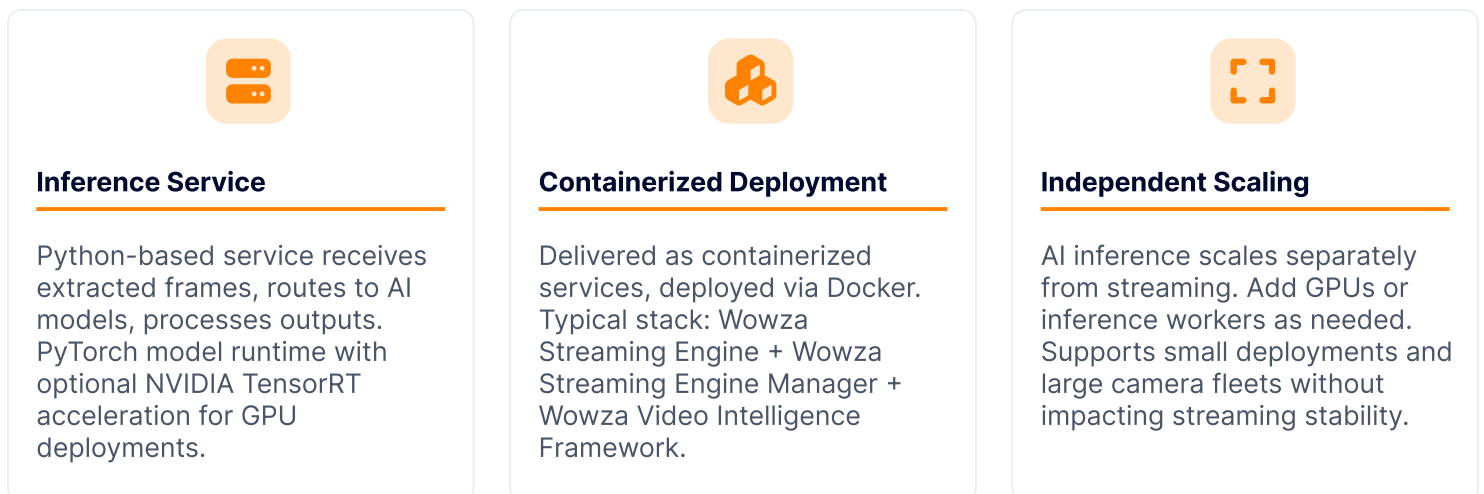
Object Detection

RF-DETR transformer-based models that identify people, vehicles, or equipment within frames. Returns bounding boxes, confidence scores, and timestamps.

Model Flexibility

Integrate your own pre-trained computer vision models. Once integrated, your model runs against live streams and produces the same structured outputs and event signals used across the platform.

DEPLOYMENT & RUNTIME ARCHITECTURE

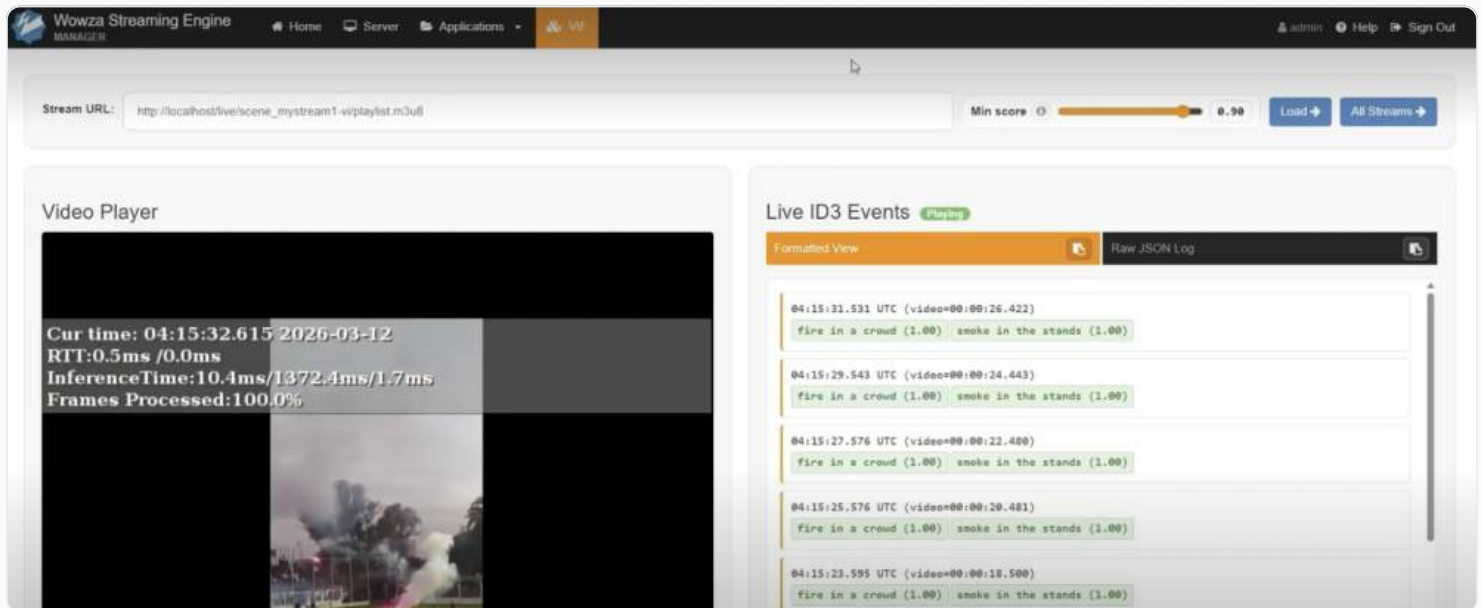


KEY CAPABILITIES

Feature	Detail
Frame Extraction	Configurable sampling intervals from live streams; no impact on the video delivery pipeline
AI Model Integration	Routes extracted frames to external models for inference; manages frame delivery, orchestration, and result handling
Model Support	Flexibility to integrate your own pre-trained models, native support for RF-DETR models
Event Signal Output	Structured metadata, thumbnails, short video clips, and webhook payloads generated from inference results
On-Premises Inference	Fully air-gapped operation supported; no cloud connectivity required during inference
Operational Integration	Webhooks, SIEM tools, incident management platforms, and traffic management systems
GPU Acceleration	Optional NVIDIA TensorRT acceleration; requires SM 7.5+ (Turing architecture or newer)

WOWZA VIDEO INTELLIGENCE FRAMEWORK IN ACTION

Live video stream processed in real time, with detections converted into structured events, alerts, and logs that teams can act on immediately.



The screenshot displays the Wowza Streaming Engine Video Intelligence Framework interface. At the top, the 'Stream URL' is set to `http://localhost/live/scene_mystream1-viplaylist.m3u8`. A 'Min score' slider is set to 0.90. The interface is divided into two main sections:

- Video Player:** Shows a live video stream of a crowd. A text overlay in the top left corner provides performance metrics:
 - Cur time: 04:15:32.615 2026-03-12
 - RTT: 0.5ms / 0.0ms
 - Inference Time: 10.4ms / 1372.4ms / 1.7ms
 - Frames Processed: 100.0%
- Live ID3 Events:** A panel titled 'Live ID3 Events' with a 'Playing' status. It shows a list of detected events in a 'Formatted View' table:

Timestamp (UTC)	Video Time	Event 1	Event 2
04:15:31.531	(video=00:00:26.422)	fire in a crowd (1.00)	smoke in the stands (1.00)
04:15:29.543	(video=00:00:24.443)	fire in a crowd (1.00)	smoke in the stands (1.00)
04:15:27.576	(video=00:00:22.400)	fire in a crowd (1.00)	smoke in the stands (1.00)
04:15:25.576	(video=00:00:20.481)	fire in a crowd (1.00)	smoke in the stands (1.00)
04:15:23.595	(video=00:00:18.500)	fire in a crowd (1.00)	smoke in the stands (1.00)

