

Intelligent Video Analytics — Project Scoping Checklist

Size any retail, industrial, robotics, or surveillance vision project before the next vendor call.

1. Define the rule — what is the one question the camera must answer?

Retail	Industrial	Robotics	Autonomous vehicle
Did they pay? Shrink, scan-avoidance, shelf gaps, shopper heatmaps.	Is the part defective? Anomaly detection on good-only data. 95-99% accuracy target.	Where am I / what is that? Detection + pose + SLAM. Camera + LiDAR fusion.	Is the path safe? Fusion or vision-only. Hard real-time, recall in bad weather.

2. The four blocks — confirm a choice for each

#	Block	What it does	Production choices
1	Detection	Box + label + score per object	YOLO family · Grounding DINO (open-vocab)
2	Tracking	Stable identity across frames	ByteTrack · DeepSORT · OC-SORT
3	Rules & analytics	Tracks become business meaning	Line crossing · dwell · scan-mismatch · anomaly
4	Deployment topology	Where blocks 1-3 run	Edge chip · central GPU · hybrid (most common)

3. Edge or cloud — tick what your deployment needs

- ☐ **Lowest latency** — alert in milliseconds, no network round-trip → edge.
- ☐ **Privacy first** — raw video must never leave the device → edge.
- ☐ **Tight bandwidth** — ship alerts, not full video streams → edge.
- ☐ **Largest model** — need a heavier or VLM model for accuracy → cloud.
- ☐ **Fleet updates** — push one model to every camera instantly → cloud.
- ☐ **Pooled GPUs** — share expensive hardware across many cameras → cloud.

4. The four traps that kill the budget

Trusting per-frame accuracy

A 95%-accurate detector over a 900-frame session fires false alerts constantly.

Fix: aggregate evidence across the whole track before alerting.

Watching only the till

~80% of stolen items are concealed in the aisle, before checkout.

Fix: aisle-to-checkout tracking, alert at the exit.

Classifying defects you can't collect

Good factories produce too few defects to train a classifier per type.

Fix: anomaly detection trained on good-only images.

Ignoring the regulation

Cameras that identify people trigger GDPR Art. 9 + EU AI Act high-risk rules.

Fix: anonymised-by-design tracking; strict retention windows.



GO DEEPER · THE 7-VOLUME HANDBOOK

Fora Soft Video Engineering Handbook

Seven volumes, 9,000+ pages: encoding, streaming, audio and AI for video products — from the pixel to production.

By Nikolay Sapunov, CEO at Fora Soft · Kindle & paperback on Amazon

[Browse the handbook on Amazon →](#)

This checklist scopes a vision project: the one rule, the four blocks, edge or cloud, and the budget traps. Our seven-volume handbook covers the whole video stack; start with Volume 6 for detection, tracking and segmentation, then Volume 7 for real-time pipelines, serving and compliance such as the EU AI Act.

WHAT YOU WILL FIND INSIDE

- Detection, tracking, segmentation and OCR (Vol. 6)
- Serving, cost per stream-hour and EU AI Act rules (Vol. 7)
- Deployment topologies: edge, central GPU, hybrid (Vol. 6)

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Scoping a retail, industrial or surveillance analytics project?

We will turn your rule into a pipeline spec with model choices, edge or cloud placement and a cost estimate per camera. 250+ video products shipped since 2005.

[Book a 30-min call →](#)

[WhatsApp →](#)

[Email us →](#)