NVA (Nota Vision Agent)

NVA (Nota Vision Agent): A Generative AI-Powered Vision Agent

NVA (Nota Vision Agent) is a Vision-Language Model (VLM)-powered surveillance solution that autonomously perceives and understands on-site situations.

It instantly detects and analyzes potential hazards from real-time video streams, dramatically enhancing safety management efficiency across industries.

Key Features

Real-time Monitoring with Instant Alerts

- Accurately detects critical events from vast image and video datasets
- Delivers instant alerts to enable swift responses to hazards and prevent accidents
- Customizable alert thresholds to meet specific operational needs

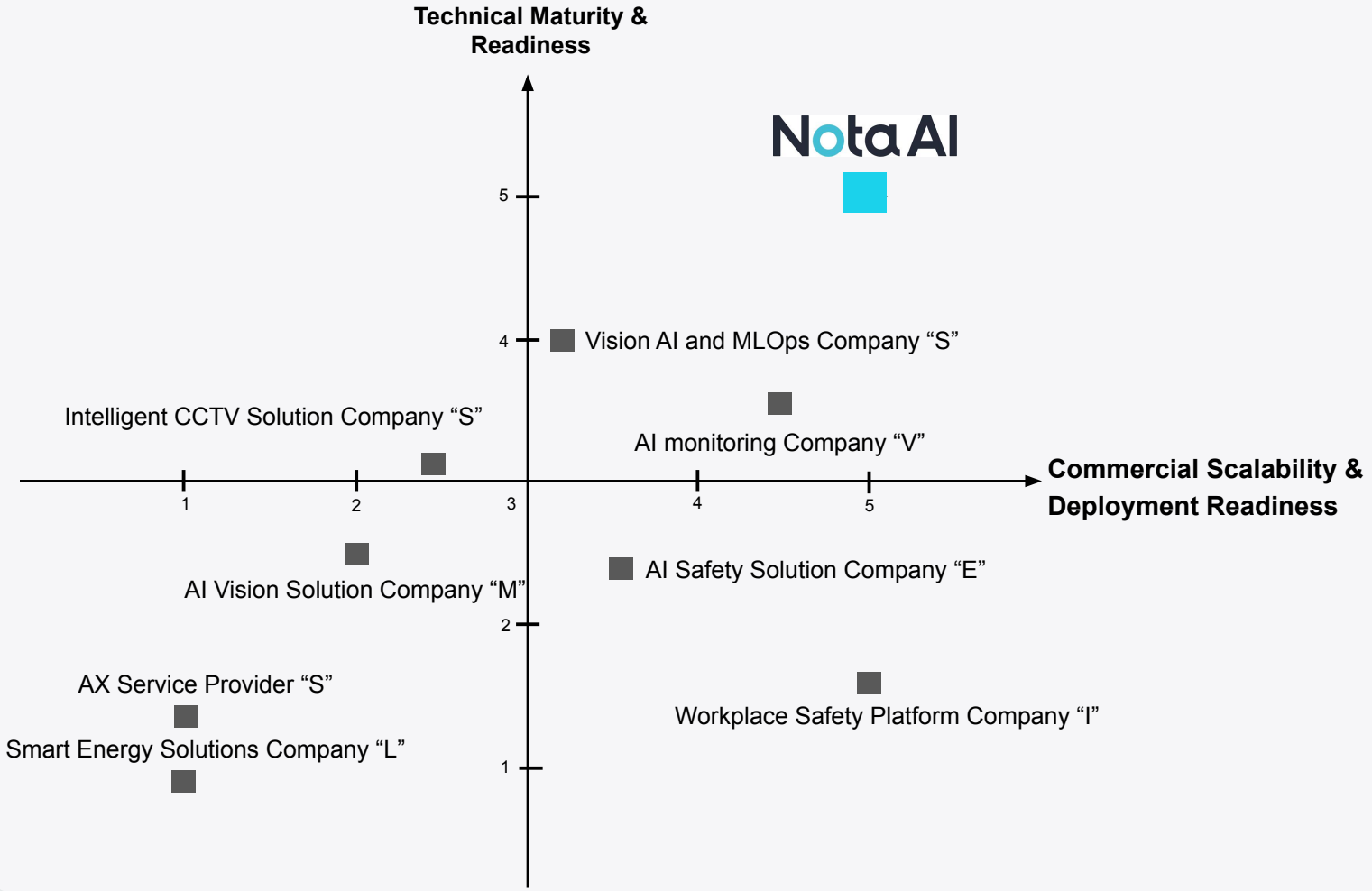
Event Summaries and Analytical Reports

- Automatically generates analytical reports by identifying and assessing event occurrences
- Provides detailed insights into key indicators and trends based on monthly or weekly event data

Natural Language Q&A and Video Search

- Analyzes large-scale visual data—from historical archives to real-time streams—to provide precise answers to user queries in natural language
- Understands contextual relationships in complex scenarios, such as sequential condition violations, enabling logical reasoning and informed decision-making

Competitive Advantage



Multi-Model Vision AI Expertise

Ability to develop and deploy vision AI models—including VLM and CV—in the most effective combination for each scenario.

Edge Deployment Capability

Designed to operate reliably within constrained hardware and network environments, powered by proprietary AI compression and optimization technologies.

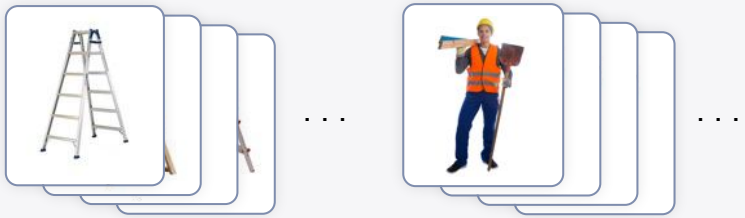
Proven End-to-End Delivery Experience

Validated through commercial deployments, scaling from PoC to full production systems.

Technical Capabilities | Differences Between CV and VLMs

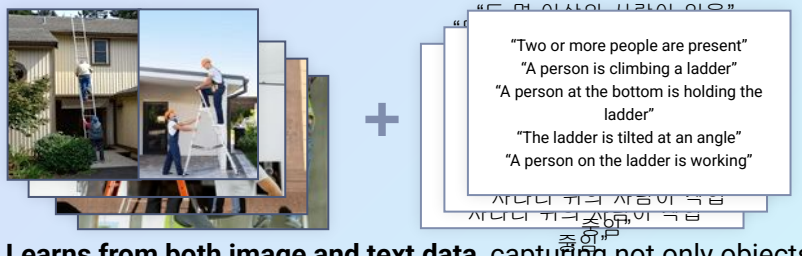
Qualitative Differences in Training Data

CV



Primarily focuses on **identifying and locating specific objects within an image**.

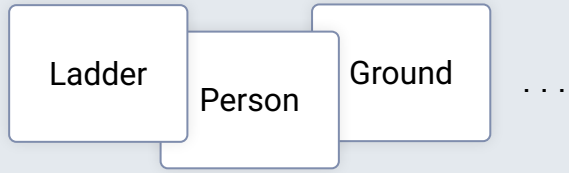
VLM



Learns from both image and text data, capturing not only objects but also contextual information such as human actions, emotions, and locations.

Example Input Image

Detection Results



- ✔ **Detects only trained objects**
- ✔ **limited understanding of surrounding context**

The worker below has turned away from the ladder and stepped away, leaving the worker above working alone at height — a violation of safety regulations

- ✔ **Goes beyond simple object recognition to understand and describe the full meaning and context of an image**

Technical Capabilities | Scenario-Based Model Application

#1

Well-defined object detection

- Detecting hazardous elements with distinct characteristic in specific areas
  - E.g. detecting localized flames within a specific area

#2

Requiring value judgement

- Analyzing and predicting potential damage levels or spread areas based on current condition
  - E.g. severity analysis based on fire size and potential spread

#3

Behavior-based Hazard Detection

- Identifying potential risks such as SOP violations, negligence, and unsafe actions to proactively prevent incidents
  - E.g. operating without safety equipment (exhaust systems or gas detectors) activated; Improper handling or storage of hazardous materials

Both **CV** and **VLM** appropriate

Only VLM

Nota AI is proficient in both CV and VLM technologies, applying the most suitable approach based on client needs and environmental condition.

Customer Benefit

Achieving High Detection Performance through Versatile Model Combination

- Combines various vision AI technologies including VLM, Visual Grounding, and CV to **address a wide range of detection requirements**
- **Customizes model configurations** based on site-specific characteristics and detection complexity

Maximizing Business Productivity through Easy Installation and Deployment

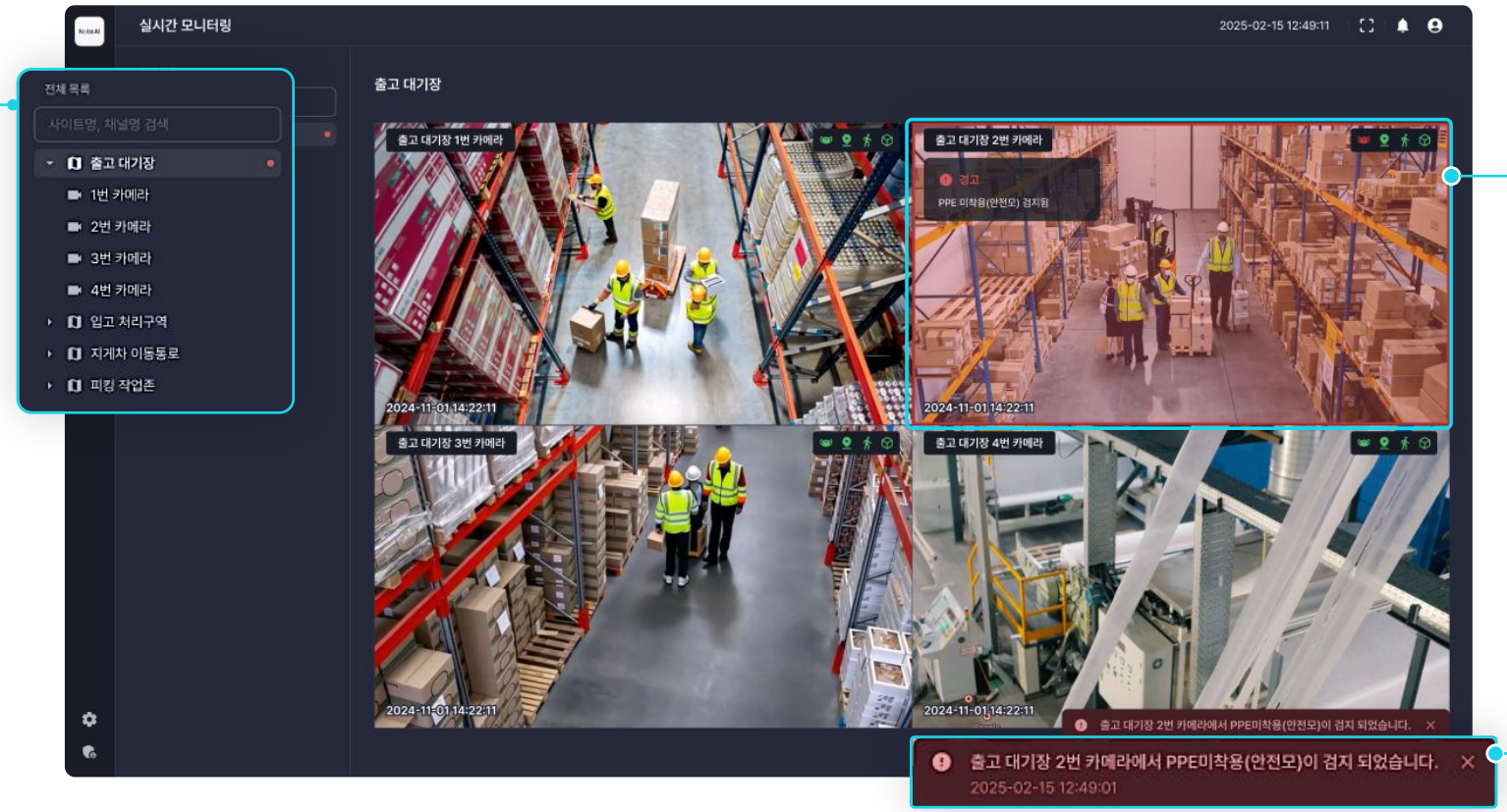
- **Compatible with existing CCTV and video management systems** without the need for additional equipment purchases
- **Can be applied on-site within 2–3 weeks without complex pre-training**, with support for channel expansion according to operational environments

Improving Operational Efficiency

- **Reduces administrative workload** with features such as automated analysis reports and video search, enhancing operational efficiency
- **Enables efficient personnel allocation** through automated CCTV monitoring

NVA In Action | Real-Time Monitoring

Searchable List of Hazard Detection Zone and Channels



Real-Time Hazard Alerts

- Identifying the zone and type of incident
- With the integrated alert system, proactive preventative measure could be implemented
- Enables quick response by monitoring personnel

NVA In Action | Dashboard

Industrial hazards are categorized, tracked, and converted into data

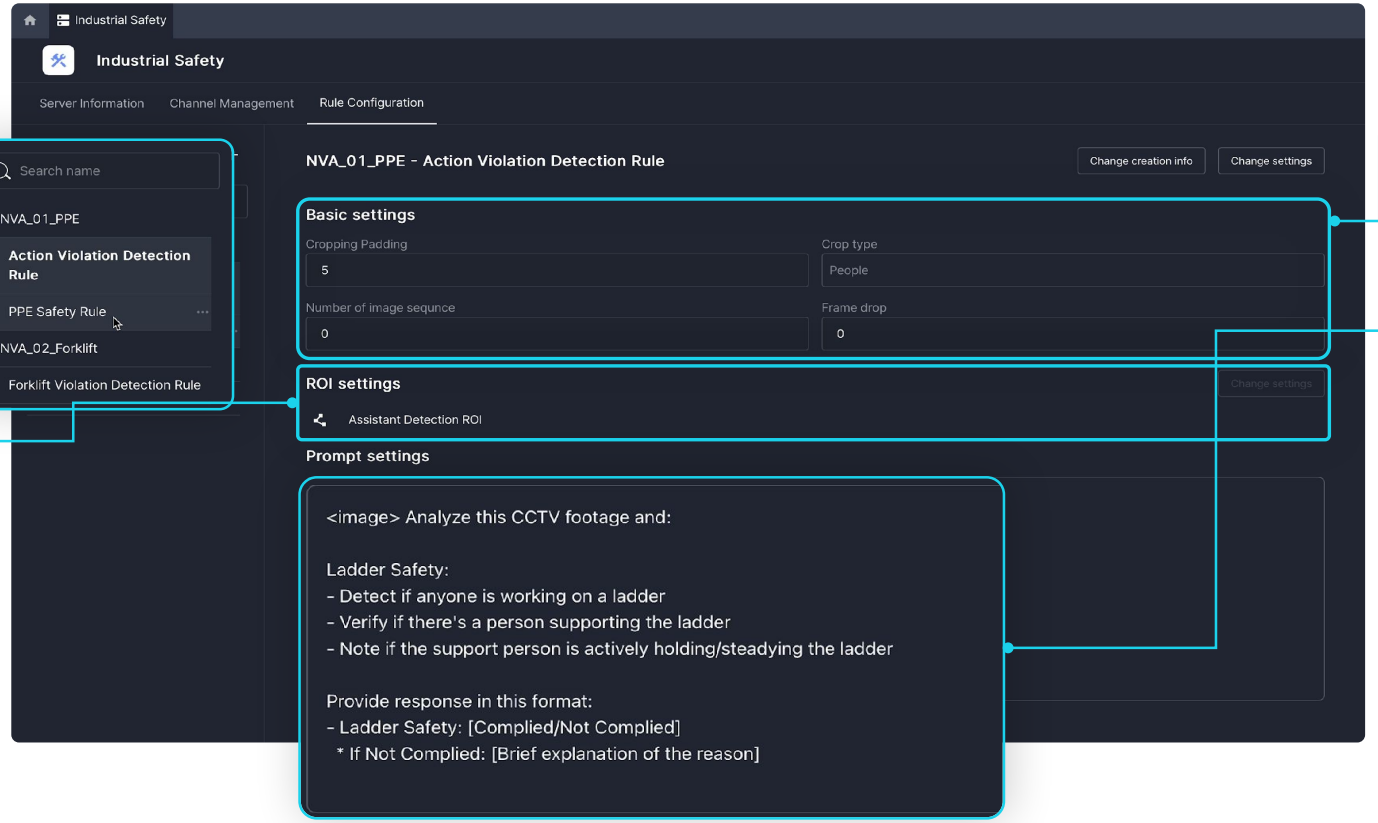
Displays the precise location of each hazardous incidents

Visualizes and analyzes incident occurrence data by month and category

NVA In Action | Rule Setting

List of Detection Tasks
- Easily search existing rules

Monitoring Specific Hazard Zone
- Allows setting customizable zones for focused analysis of collected video and images
- E.g. pedestrian-only or restricted-access areas



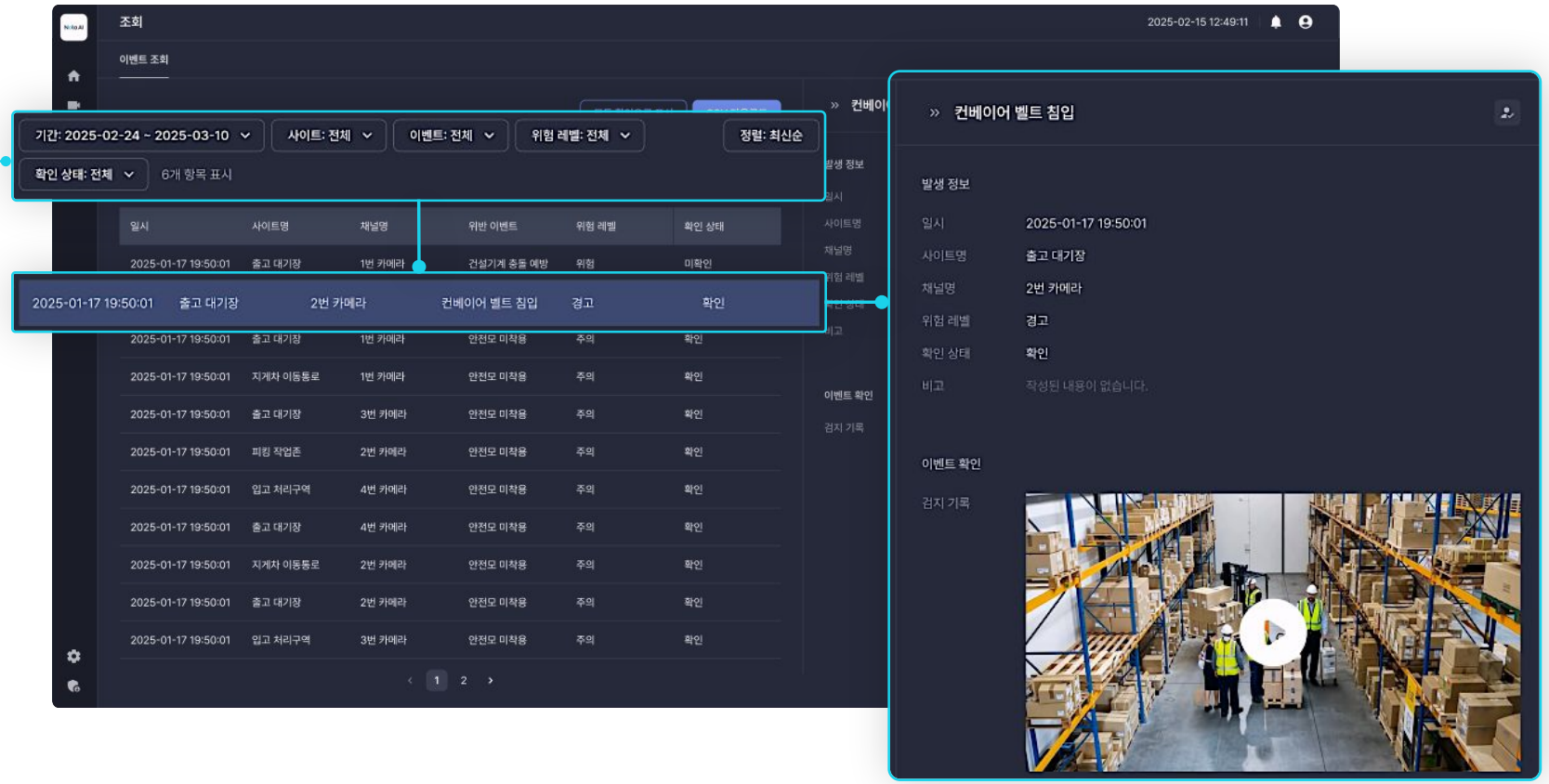
Video Analytics Setting

Add or Edit Detection Rules
- Enter hazard detection requirements in natural language prompts
- Significantly faster rule creation and application compared to traditional CV solutions
- No coding required

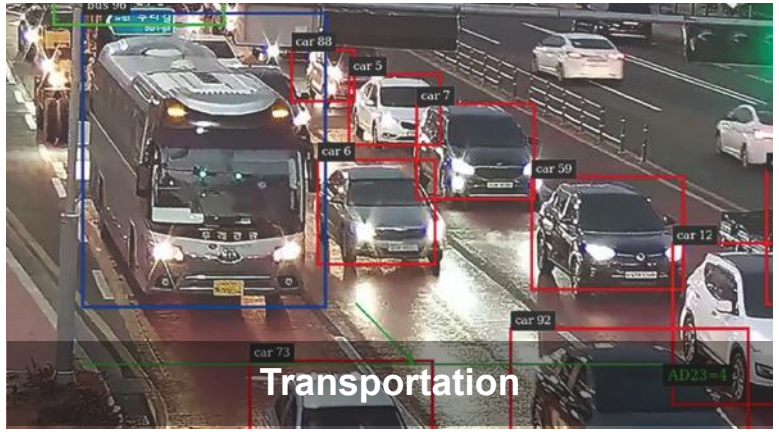
NVA In Action | Incident Search

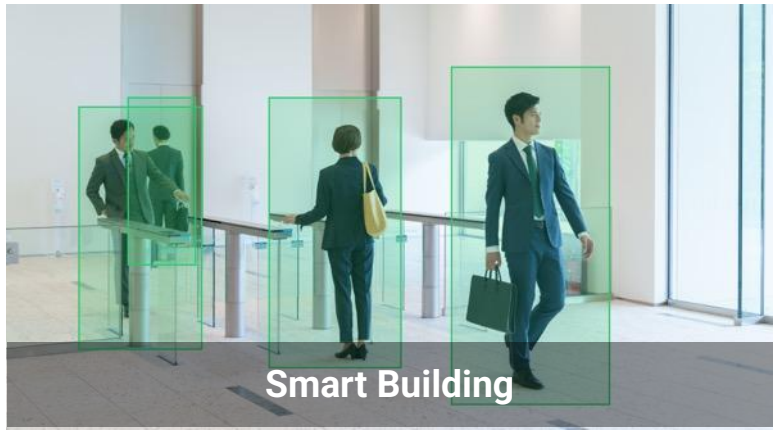
Search Specific Incident

- Search specific events by applying filters (e.g. time, type of incident, etc) from all collected footage
- Supports natural language search
- Eliminates the need for manual review of all footage



Applicable Industries





Use Cases - Industrial Safety

Chemical & Textile Manufacturer “K”

Challenge

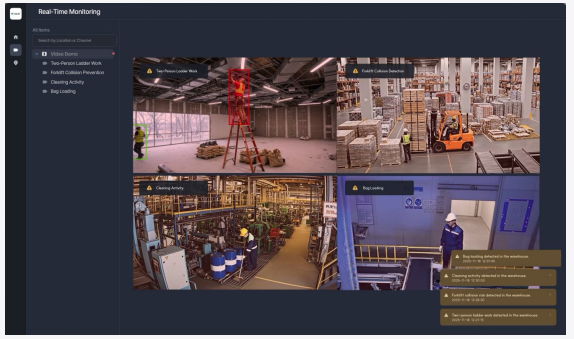
- Traditional monitoring systems lack the ability to interpret complex worker behaviors
- Personal protective equipment (PPE) violations and unsafe behaviors often go unnoticed until incidents occur

Solution

- Equipment Interlock automatically halts machinery when workers enter dangerous zones
- SOP Compliance Monitoring detects violations such as improper cleaning, incorrect loading, and floor-level repackaging through VLM-powered analysis

Result

- Proactively detects and blocks human error-based hazards, preventing accidents at the source
- Achieved F1 Score of 85+ for SOP compliance accuracy across all tested tasks



Use Cases - Surveillance

Municipal Government "G"

Challenge

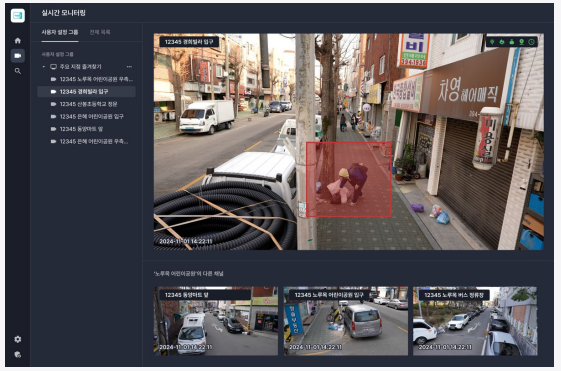
- Illegal dumping is difficult to monitor consistently across widespread public areas
- Smoke and burning activities often go undetected until fire incidents escalate

Solution

- Edge-deployed NVA detects smoke and burning activities to prevent fire incidents in real-time
- Contextual video intelligence identifies illegal dumping behaviors that traditional monitoring systems miss

Result

- Improved response efficiency for public safety incidents
- Enabled consistent monitoring while maintaining strict privacy and data governance compliance



Use Cases - Transportation

Abu Dhabi Integrated Transport Centre (ITC)

Challenge

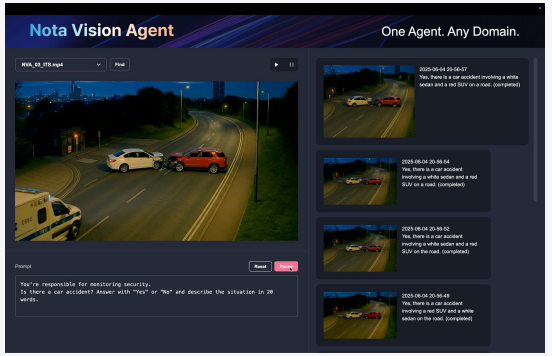
- Rapid detection of road incidents is critical across vast highway networks with high-speed traffic
- Cloud-dependent systems pose challenges in latency, operational costs, and data security compliance

Solution

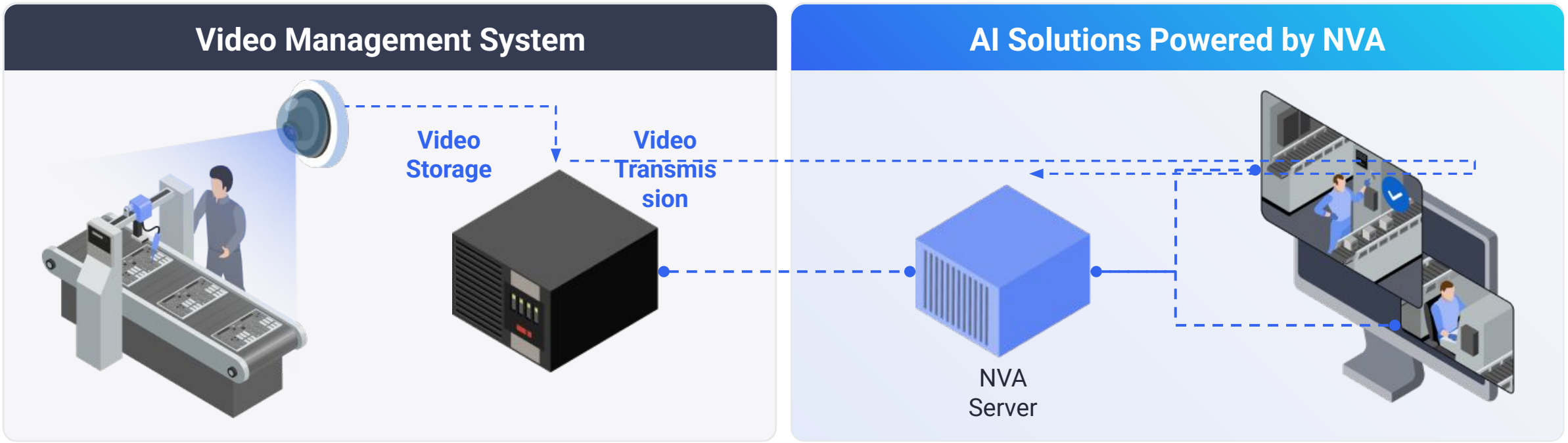
- VLM-powered NVA detects road incidents and anomalies in real-time directly on edge devices
- On-device processing ensures minimal latency while maintaining strict data privacy and security standards

Result

- Achieved 95%+ accuracy in road incident detection during PoC validation
- Reduced cloud dependency, lowering operational costs while enabling real-time response



Pipeline



- **No additional equipment required –** Easily integrate by connecting the ‘NVA server’ to existing video management systems.
- While conventional computer vision (CV) solutions require large-scale data training and take 3–6 months for deployment, the Vision-Language Model (VLM)-based **NVA can be rapidly implemented in the field within just 2–3 weeks**, without complex pre-training processes.