Elevating Safety and Efficiency in Mining with Vision AI: From Object Detection to LLM-Driven Decision Intelligence

Yale Zhang , M.A.H. Zahid , and Thessygan Moodley

Presented by Dawid Jordaan

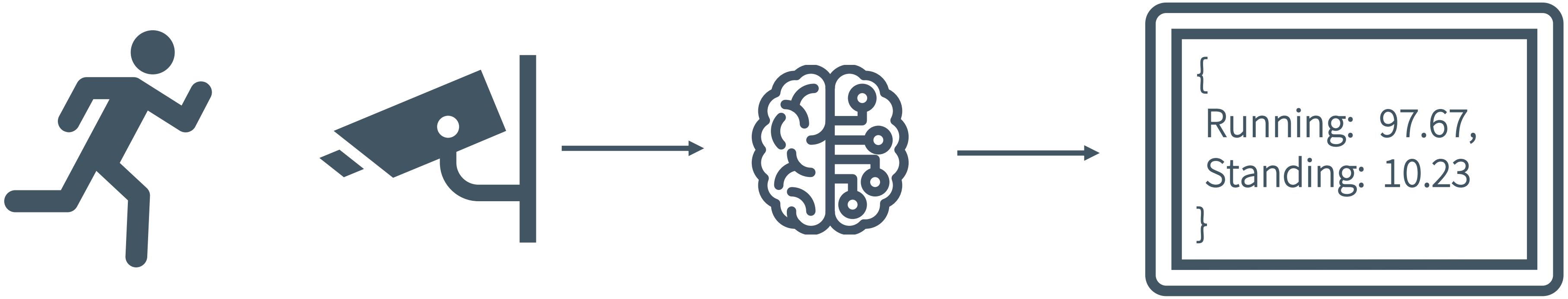
28 October 2025

HATCH

1. What Is Vision AI
2. Challenges Facing The Mining Industry
3. Traditional Vision AI Implementations
4. The Evolution To Advanced Multi-modal Integrated Vision AI Systems
5. 2 Case Studies

Vision AI: Turning Pixels into Insights

Automatic image/video understanding and analysis



Object Detection

Event Detection

Movement Detection

Position Detection

HATCH

Why Vision AI In Mining

Addressing The Gap Between What Is Visible And What Is Tracked And Responded To

+ Vision AI bridges the gap by converting images into real-time signals that can lead to actions.
+ Detects critical safety issues and operational bottlenecks.
+ Provides timely alerts and meaningful recommendations so that operators can intervene early.

Wide Applicability

Conveyor Monitoring	Crusher Oversight
Maintenance Safety	Environmental Compliance

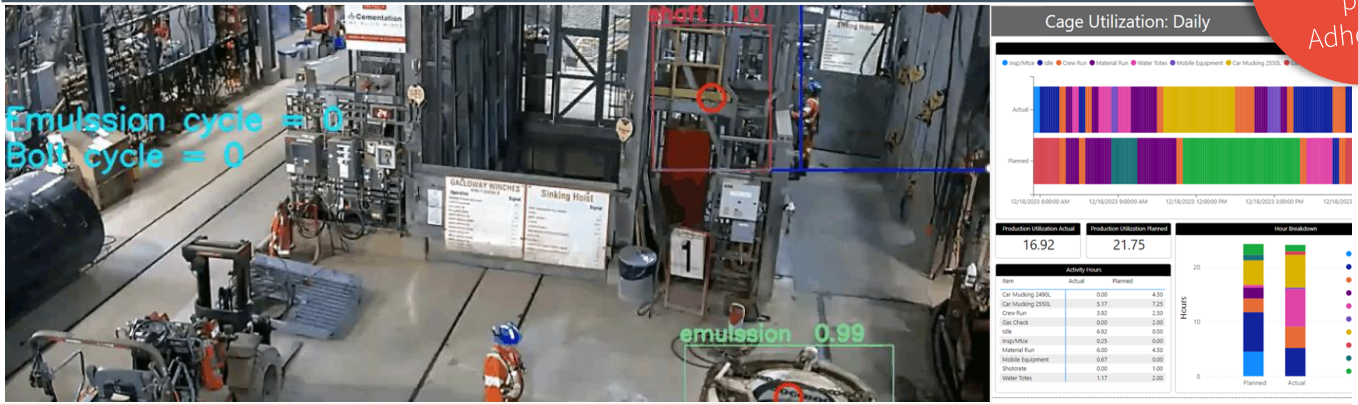
HATCH

Traditional Object Detection

Powerful Recognition, But Limited Understanding

Underground Mine Operation Monitoring and Event Tracking

3x
Production plan Adherence



Mine cage was a major operational bottleneck, with delays causing production setbacks and budget overruns as high as 30%.

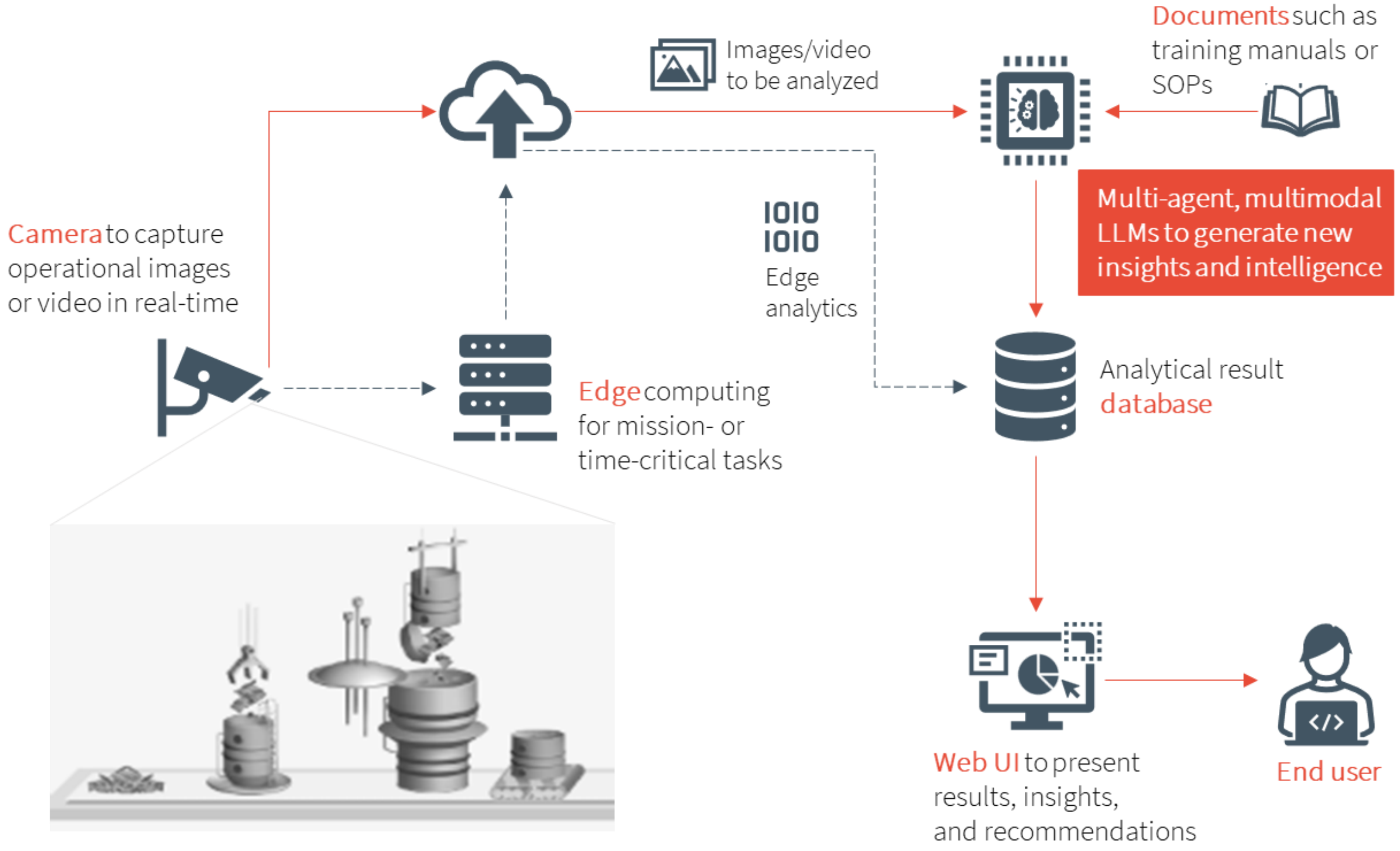
- Cameras installed at shaft stations
- YOLO models timestamped events (e.g. material runs, gas checks)
- Real-time data compared against production schedules to flag deviations and bottleneck

Technology

- Single-stage object detectors such as **YOLO** (You Only Look Once).
- Models excel at localising and identifying objects (e.g. people, equipment, vehicles) in real-time.
- Effective in structured environments with well-labelled data.
- Require large specifically labelled datasets for each environment and task.
- Requires retraining if camera angles, lighting or equipment changes.
- Lack context and understanding – needs augmentation through heuristics

LLM-driven Decision Intelligence

Visual Context Evolving To True Operational Understanding



Technology

Gemini

LLaMA by Meta

ChatGPT

- Leverage multimodal LLMs for deep, context-rich reasoning.
- Eliminate reliance on large pre-labeled datasets.
- Dramatically cut model-development and maintenance effort.
- Maintain robustness across diverse camera angles and lighting conditions.
- Introduces understanding of operational context enabling detection/description of events and objects not explicitly fine-tuned for.

HATCH

Case Study 1: Open-pit Mine Operations

Leverage Existing CCTV Streams to provide actionable insight on productivity and safety.

Objectives:

- Generate cycle analytics for trucks and shovels
- Detect unsafe or inefficient behaviours

Case Study 1: Open-pit Mine Operations

Boosting Throughput And Proactive Safety With Existing Cameras At An Open-pit Mine

Video Footage	Automated Report

** Analysis of Excavator Loading Haul Truck Footage

Key Equipment Labels:

- Excavator: CAT, 6090
- Haul Truck: CAT, 793F, 56

Analyze Each Haul Truck Activity:

- **Truck Arrival Time**: 0:00
- **Truck Positioning Time**: 0:00
- **Loading Process**
- **Number of Loads**: 3

Calculate KPI Metrics:

- **Start Time and End Time of Each Load:**
- Load 1: Start - 0:09, End - 0:14
- Load 2: Start - 0:26, End - 0:30
- Load 3: Start - 0:45, End - 0:49

The fine-tuned model achieved an **accuracy rate of approximately 98%** in delivering domain-specific analytics.

HATCH

Case Study 2: Smelter Operations

Transforming unstructured video footage to structured data to improve logistics and scheduling.

Objectives

- Identify operational safety events
- Produce time-stamped evidence linked record of events

Case Study 2: Smelter Operations

Creating A Structured, Actionable Record Of Events From Converter-aisle Video

Video Footage

Automated Report

Event 1: A scrap boat, marked with the ID "18," is positioned directly below the convertor. The scrap boat is a rectangular container designed to hold scrap metal and is suspended horizontally from cables.

Event 2: The scrap metal is charged from the scrap boat into the convertor. This is done by tilting the scrap boat and allowing the metal to pour into the furnace. The molten metal appears bright orange and sparks are seen flying. The sound of the metal pouring into the furnace is audible.

Event 3: The scrap boat, still marked "18," moves away from the convertor. The furnace roof begins to close, sealing the furnace for the smelting process.

Event 4: A ladle, marked with the ID "M3," is positioned below the convertor. The ladle is a cylindrical container designed to hold molten iron ore and is suspended vertically from cables.

Event 5: The iron ore is roasted and smelted within the convertor. This process is not directly shown in the video, but the furnace's intense glow indicates the heat of the process.

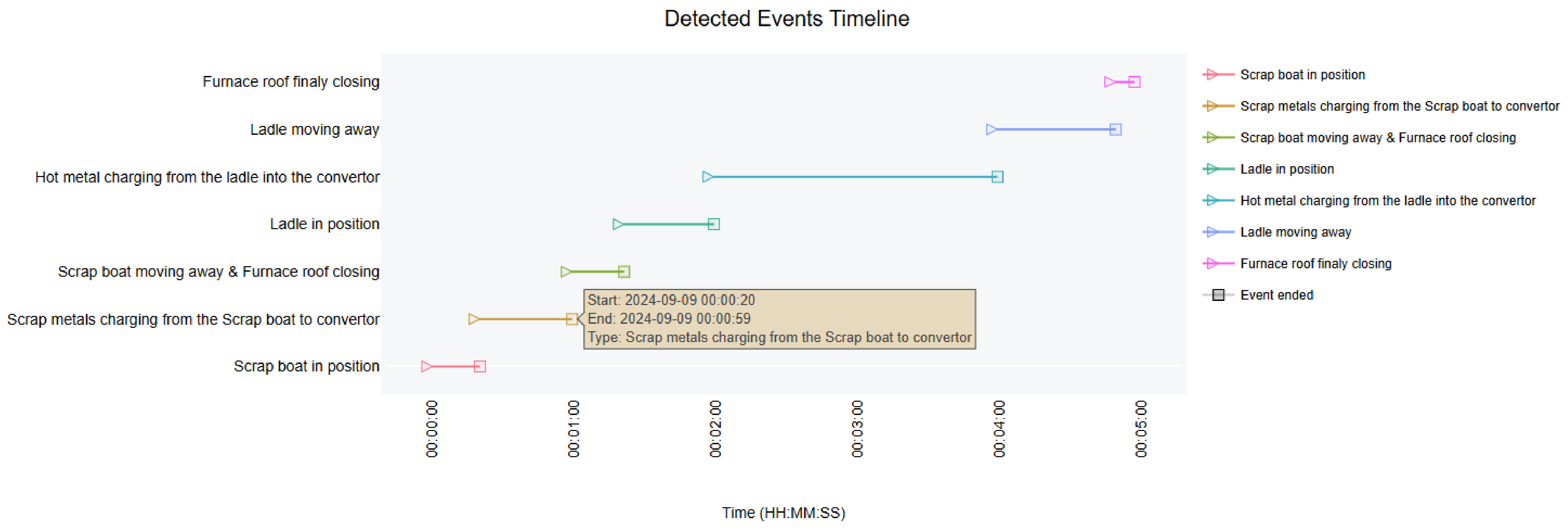
Event 6: The ladle, still marked "M3," is moved away from the convertor. It is presumably being transported to a casting station or other processing area.

The system generates a synchronized, human-readable description of the events as they unfold.

HATCH

Case Study 2: Smelter Operations

Creating A Structured, Actionable Record Of Events From Converter-aisle Video



The system creates a timeline of all key activities, with precise start and end timestamps.

Practical Considerations For Implementation

	Traditional Object Detection	Multimodal LLMs
Performance	**Strengths:** High speed with low computational cost. **Best Use Case:** Stable environments with well defined tasks requiring immediate response	**Strengths:** Superior contextual understanding, adaptability and procedure awareness. **Best Use Case:** Interpreting complex, dynamic scenes to provide structured data that is procedure aware without constant retraining.

Implementation

Ruggedised Hardware:
Systems must be built to last. Cameras and edge computers require industrial-grade specifications to withstand **extreme temperatures, dust, moisture, and vibration.**

Seamless Integration:
Value is maximised when Vision AI connects to your existing ecosystem. This includes integration with **historians, maintenance systems (CMMS), and safety reporting platforms.**

System Reliability:
System latency and accuracy should be aligned with the operational objectives to ensure timely, trustworthy notifications while avoiding false positives.

Return on Investment

Measurable Economic Gains:

- The case studies demonstrate a direct impact on the bottom line, including a **23% increase in material movement and a 31% reduction in scheduling deviations.**
- This translates to increased capacity without additional capital investment.

Invaluable Safety Improvements:

- While harder to quantify, the safety benefits are profound, including reduced incident rates and improved compliance monitoring.
- The paper makes a critical point: **The value of preventing a single serious safety incident often justifies the entire system investment**

A Shift In Vision AI Capabilities

An advancement from Recognition to Understanding

- From detecting predefined tasks to **understanding the full operational sequence** and its context.
- Multi Modal LLMs **incorporates additional knowledge from procedure and site documents** for a deeper understanding of the tasks at hand.

Adding value in complex and Dynamic Environments

- Measurable **improvements to safety and efficiency** as proven in both case studies
- Offers **superior adaptation to environment variations** (e.g. camera angles, lighting) reducing the need for continuous model retraining.

Future is Autonomous and Procedure-Aware

- A key **enabler for autonomous operational monitoring** with integrated decision making.
- The additional procedure awareness and unstructured to structured data transformations can **lead to competitive advantages in achieving performance and safety outcomes.**

HATCH