ICPR 2026 - MCMI Workshop

Detect, Anticipate, and Explain Anomalies in Agricultural Robots: A Review

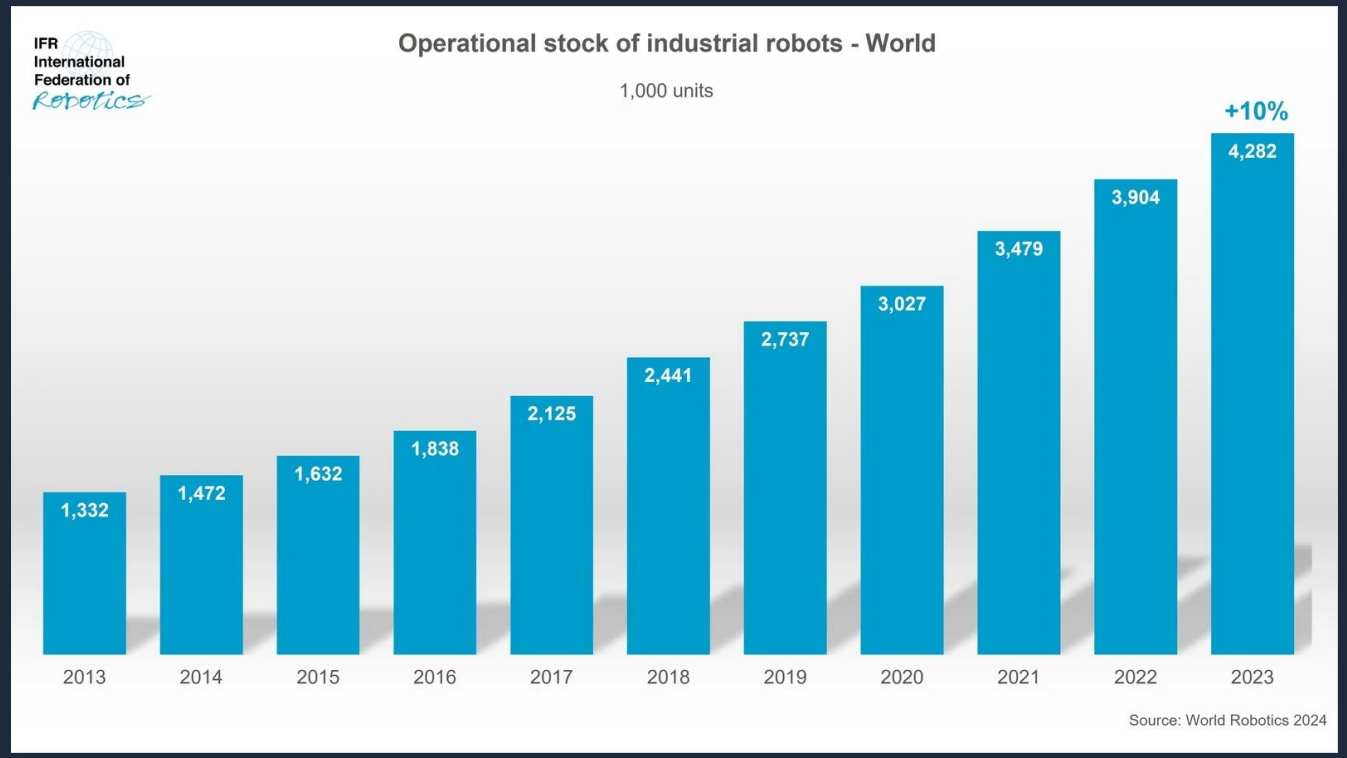
Thomas Trouvé
INRAE, Neusta
Co-authors: Damien Appert, Riccardo Bertoglio, Sandro Bimonte, Jocelyn De-Goer-De-Herve

INRAE

neusta france

1. Context

- Robots are increasingly deployed beyond controlled environments
- Agriculture is one of the major domains for autonomous robotics
- Unlike industrial robots, agricultural robots face unpredictable environments



INRAE

neusta france

2. Motivation

- Unexpected anomalies are inevitable during field operation

Risk of :

- Operational failure
- Hardware damage
- Downtime

Need for anomaly management:

- Detect
- Anticipate
- Explain

INRAE

neusta france

3. Four Core Challenges in Smart Farming

Data Scarcity

Rare, diverse, and costly-to-collect failures

-> Unsupervised Learning

Multimodal Data

Agricultural Robots rely on heterogeneous sensors

(camera, GPS, IMU)

-> Multimodal Fusion

Ressources & Connectivity

Limited onboard resources and unreliable connectivity

-> Edge AI

Non-Expert Operators

Technical outputs are difficult to interpret

-> Clear, Interactive and accessible explanations

INRAE

neusta france

4. What is multimodality in Robotics?

Human Perception

Different ways of perceiving information

- Vision
- Hearing
- Touch
- ...

Machine learning

Different data representations

- 2D Image
- 1D Signal
- Tokens
- ...

Robotic systems

Multimodality is used inconsistently

- Often confused with multiples sensors
- Multi-sensor ≠ Multimodal
- Multimodal data ≠ multimodal problem & solution

INRAE

neusta france

5. Proposed Framework

Multimodal Data

Presence of fundamentally incompatible structures.

- Structurally incompatible, cannot be merged in their raw state.

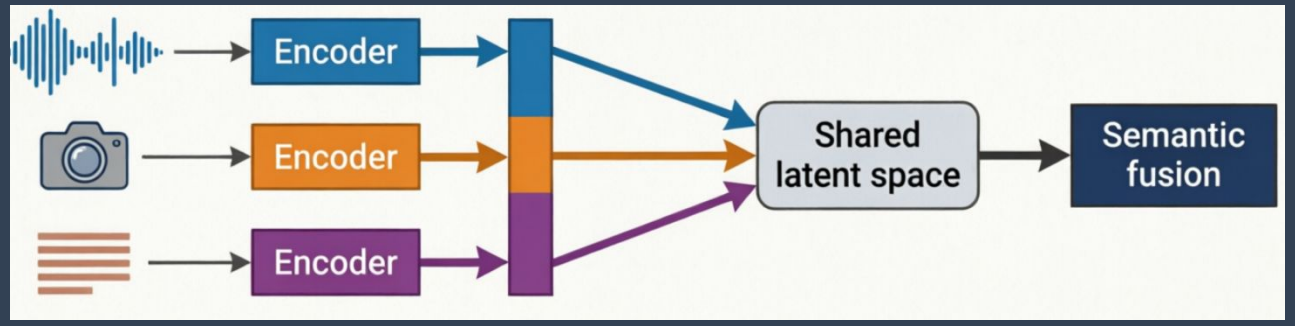
Problem Multimodality

The task requires heterogeneous inputs to be solved

- Not Multimodal: Obstacle detection
- Multimodal: Instruction-guided manipulation

Solution Multimodality

Dedicated modality-specific transformations.



INRAE

neusta france

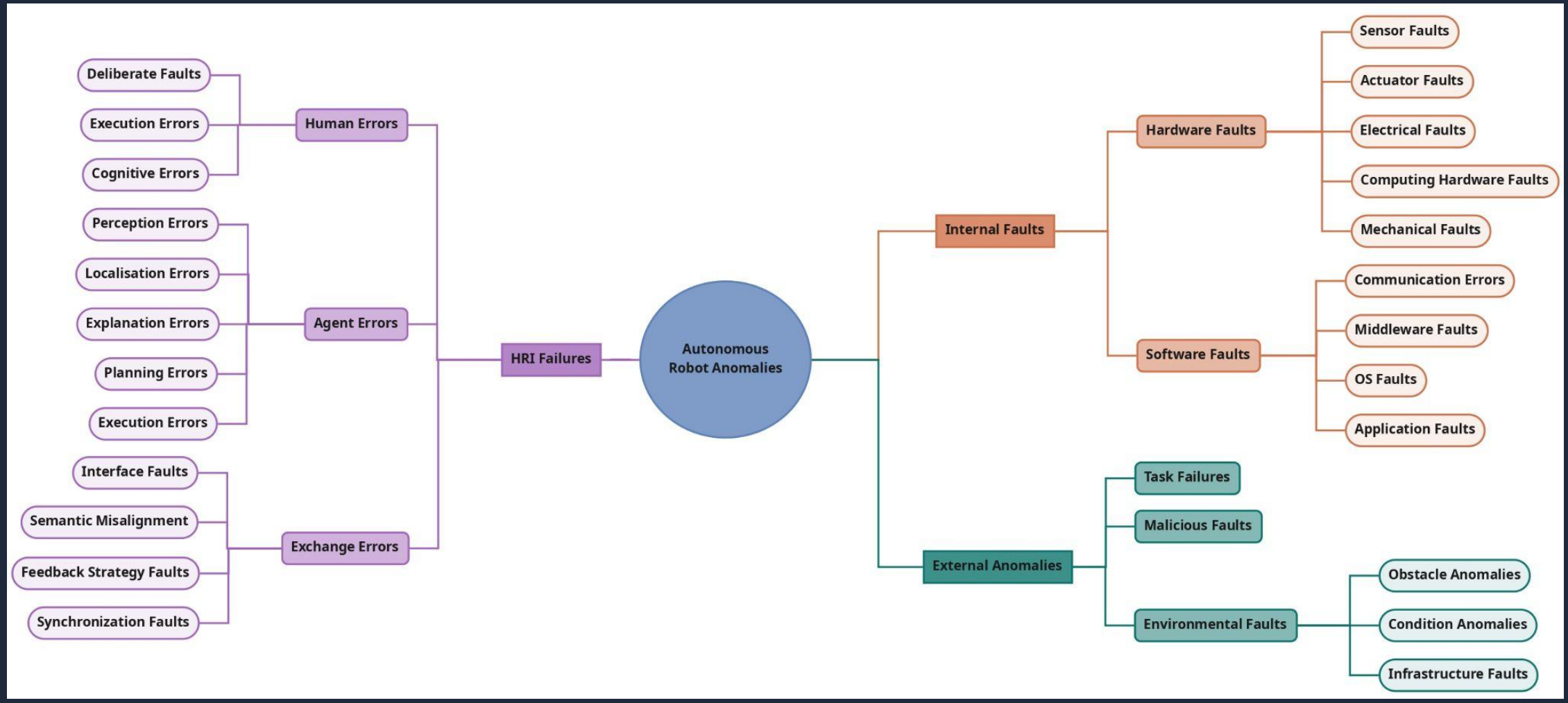
6. Taxonomy of Robot Anomalies

Internal	External	HRI Anomalies
Hardware Fault	Task Failures	Human Errors
Software Fault	Malicious Faults	Agents Errors
	Environmental Faults	Exchange Errors

INRAE

neusta france

7. A Novel Taxonomy of Robot Anomalies



8. Summary of Reviewed Studies

Article	D	A	E	Sup.	Inputs	Topology	Model	Anomaly Scope
37	✓	✓	×	P	R, T	2D, Tk	CLIP + DINOv2	x(env, task), hri(human, agent)
39	✓	✓	✓	P	Pr, R, T	1D, 2D, Tk	GMR + CondNet	i(hw), x(task), hri(human, agent)
29	✓	×	✓	F	T, R, Dp, Au, Pr	1D, 3D, Tk	CLIP + GPT-4	x(env, task), hri(agent)
36	✓	✓	✓	F	R, T	2D, Tk	GPT-4V	x(task)
20	✓	×	✓	F	T, R, L	2D, 3D, Tk	GPT-4	x(task)
42	✓	×	×	F	R, Dp, Au, Pr	1D, 2D	CNN + DAE	x(task, env), i(hw)
26	✓	×	✓	F	R, T	2D, Tk	GPT-4o + QwenVL	x(task)
23	✓	✓	✓	F	T, R, L, I, G	1D, 2D, 3D, Tk	Llama2-13B	x(env), hri(agent, exch.)
35	✓	×	×	F	R, Pr	1D, 2D	Autoencoder	x(env), i(hardware)

9. Modalities & Sensors

- **RGB Cameras:** The foundational modality, appearing in 100% of the reviewed studies.
- **Text Inputs:** 2nd most common. Driven by LLMs acting as semantic bridges.
- **Proprioception & Depth:** Frequently used in manipulation tasks but underutilized globally.
- **Navigation Sensors:** LiDAR, IMU, and GPS used fot navigation tasks.
- **Navigation Sensors:** LiDAR, IMU, and GPS used fot navigation tasks.

10. Application Domains

Object Manipulation

Most explored domain. High performance in controlled environments, but highly **task-specific**.

Autonomous Navigation

Real-time reactive planning. Operates in less controlled settings, closer to **agricultural constraints**.

General Monitoring

Summarizes robot experiences. Clear trend towards versatile, **multi-task frameworks**.

11. Identified Research Gaps

Reactive Systems

Reactive anomaly management

Detection after failure

Explainability

Moving from binary alerts to rich, **natural language explanations** and automated correction plans.

Physics Blindness

Limited understanding of robot dynamics

INRAE

neusta france

12. Future Research Directions

Two-Speed Architectures

- Fast: Real-time detection.
- Slow: Contextual explanation.

Small Edge Agents

Offline, lightweight natural-language assistance.

World Models

Simulating environments for proactive forecasting.

INRAE

neusta france

Thank You for Your Attention

Thomas Trouvé

MCMI Workshop - 2026

INRAE

neusta france