Scientific Research Publishing

Open Journal of Applied Sciences, 2026, 16(7), 2587-2603
https://www.scirp.org/journal/ojapps
ISSN Online: 2165-3925
ISSN Print: 2165-3917

Inverse Detection for Deterministic Object Selectivity in Industrial Safety Monitoring

Gilles Verschueren, Robbe Noens, Ward Nica, Marc Juwet, Dino Accoto

Department of Mechanical Engineering, KU Leuven, Gent, Belgium
Email: gilles.verschueren@kuleuven.be

How to cite this paper: Verschueren, G., Noens, R., Nica, W., Juwet, M. and Accoto, D. (2026) Inverse Detection for Deterministic Object Selectivity in Industrial Safety Monitoring. *Open Journal of Applied Sciences*, **16**, 2587-2603.
https://doi.org/10.4236/ojapps.2026.167144

Received: June 5, 2026
Accepted: July 26, 2026
Published: July 29, 2026

Abstract

Industrial safety sensors commonly detect object presence in monitored zones around hazardous machinery, but they generally do not identify whether the detected object is a human, product, fixed structure, or mobile system. Treating all detections conservatively is necessary for safety, but can cause unnecessary speed reductions or protective stops when known non-human objects frequently enter monitored zones. This paper proposes inverse detection, a deterministic object-filtering principle for improving selectivity in zone-based industrial safety monitoring. All detections are treated as human-equivalent by default. Detections may only be removed from the human-equivalent set when they can be verified as known non-human objects using explicit information such as geometry, pose, communication status, sensor consistency, and timeout conditions. Failed or incomplete verification returns the detection to the safety-zone evaluation. The framework consists of detection, filter-definition, verification, and safety-zone evaluation layers. Verification levels are defined for fixed structures, robot-manipulated objects, communicating mobile systems, and human-equivalent detections. A proof-of-concept implementation with multiple 2D LiDAR sensors shows that verified packages and mobile objects can be filtered, while humans, unknown objects, displaced objects, and communication failures remain active in the safety evaluation. The implementation is not a certified safety function, but demonstrates the feasibility of deterministic filtering as a transparent alternative to direct human classification.

Keywords

Inverse Detection, Deterministic Filtering, Object Selectivity, Industrial Safety Monitoring, Human-Robot Collaboration, Functional Safety

1. Introduction

Industrial robots are increasingly used in workspaces shared with human opera-

tors, transported products, and mobile systems. In such environments, safety monitoring is commonly based on object presence in predefined zones. Depending on the selected safety concept, the zone state may maintain normal operation, reduce speed, or initiate a protective stop. Examples include fixed protective fields based on ISO 13855 [1] and speed-dependent zones used in Speed and Separation Monitoring (SSM) [2].

A fundamental limitation of many industrial safety-sensing systems is that they provide geometric presence information, but not semantic object identity. A safety scanner or other electro-sensitive protective equipment (ESPE) can indicate that an object has entered a monitored zone, but it does not generally distinguish between a human operator, a product, a fixed structure, or an automated guided vehicle (AGV). Consequently, all detections must be treated conservatively as potentially safety-relevant. This behavior is necessary for safety, but it can reduce productivity when known non-human objects frequently enter monitored zones during normal operation.

Object classification appears to offer a direct solution. Deep-learning approaches enable human detection using cameras [3], LiDAR [4] [5], radar [6] [7], and multi-sensor fusion [8]. A broader overview of sensing technologies for human-robot collaboration is provided by [9], who highlight the popularity of structured-light cameras due to their ease of deployment. AI-based systems can also estimate human posture, trajectory, and intention [10] [11], which may support more selective or adaptive safety monitoring.

However, classification-driven perception introduces failure modes that are difficult to reconcile with safety-related decision making. In zone-based monitoring, a false stop caused by classifying a non-human object as human mainly affects productivity. A missed human detection or incorrect human-to-object classification is more critical, because it may prevent the required safety response. Relevant limitations include:

- **Missed detections:** a human may remain undetected.
- **Incorrect classifications:** a human may be classified as a non-human object, or vice versa.
- **Dependence on training data:** performance depends on the similarity between the deployment environment and the training distribution.
- **Sensitivity to rare conditions:** lighting changes, reflections, occlusions, unusual poses, and unseen objects may degrade performance.
- **Probabilistic inference:** neural networks output confidence scores rather than deterministic guarantees.
- **Black-box behavior:** verification and validation remain challenging. Although explainable-AI methods can support interpretation, saliency maps and attention mechanisms are not sufficiently consistent as the sole basis for certified safety decisions [12].
- **Safety-standard compliance:** many research prototypes in human-robot collaboration do not apply appropriate safety protocols [13].

These limitations matter because functional-safety standards require very low dangerous failure rates and validated diagnostic behavior [14]. For example, Performance Level d (PL d) under ISO 13849-1 corresponds to a probability of dangerous failure per hour, PFH_d, in the order of 10^{-7} to 10^{-6}. A perception model with 99.9% recall still misses approximately one object in every 1000 labelled instances. This comparison is illustrative rather than quantitative: recall is not directly equivalent to PFH_d or to an operational dangerous failure rate. It nevertheless highlights the mismatch between statistical perception metrics and safety-function reliability measures [15].

ISO/IEC TR 5469:2024 therefore treats AI in functional safety cautiously and indicates that AI-based functions should not be relied upon as the sole safety mechanism without additional validated safeguards [16]. Several studies combine AI perception with certified safety scanners as fallback mechanisms [17] [18]. Nevertheless, certification pathways for AI-driven perception remain challenging and are not yet generally established [19]. Redundant AI-based sensing architectures, such as systems using multiple AI-trained cameras [20], may improve robustness but increase cost, complexity, and validation effort.

This creates the central trade-off addressed in this paper: classification-driven perception can improve object selectivity, but remains difficult to validate or certify as a stand-alone safety mechanism; deterministic safety sensing is easier to validate for presence detection, but typically lacks object selectivity.

This paper proposes *inverse detection* as a deterministic object-filtering principle for this trade-off. Instead of attempting to prove that a detected object is human, every detection is treated as human-equivalent by default. Detections associated with known non-human objects may only be removed from the human-equivalent detection set after deterministic verification based on explicit information such as geometry, pose, communication status, robot state, sensor consistency, and timeout conditions. If verification is incomplete, inconsistent, or unavailable, the detection remains human-equivalent and is forwarded to the safety-zone evaluation layer.

Inverse detection is related to deterministic safety concepts such as muting, blanking, and dynamic protective-field switching, but differs in how selectivity is applied. Conventional muting or blanking typically suppresses predefined parts of a safety function under specific material-flow or machine-state conditions. Dynamic protective fields adapt the active safety zones themselves, usually by switching between predefined field sets or monitoring cases. In contrast, inverse detection does not define or switch the protective fields. It operates before the safety-zone evaluation and determines which detected points remain human-equivalent. Multiple known non-human objects can therefore be filtered. Dynamic protective fields and inverse detection are not mutually exclusive: the fields may define the active zones, while inverse detection determines which verified non-human detections are excluded from their evaluation.

The contribution of this paper is threefold. First, it introduces inverse detection

as a deterministic alternative to classification-driven object selectivity. Second, it presents a layered framework based on detection, filter definition, verification, and safety-zone evaluation. Third, it evaluates the concept in an industrial logistics scenario involving robot-manipulated packages, a communicating mobile object, and human intrusions. The implementation is not presented as a certified safety function, but as a proof of concept demonstrating the feasibility and limitations of the proposed filtering logic.

2. Inverse Detection Principle

2.1. Definition

Inverse detection is a deterministic object-filtering principle for improving selectivity in safety-monitored industrial workspaces. The concept reverses conventional classification logic. Instead of asking whether a detected object is human, the system asks whether the detection can be verified as belonging to a known non-human object.

All detected objects are treated as human-equivalent unless they can be deterministically verified as known non-human objects eligible for filtering.

Here, *human-equivalent* means that a detection is treated conservatively as potential human presence during safety-zone evaluation. A human-equivalent detection can therefore trigger reduced speed or a protective stop. The term *deterministic* means that filtering follows explicit geometric, temporal, and communication-based rules.

A detection may only be removed from the human-equivalent detection set if all required verification conditions are satisfied. These conditions may include known object geometry, expected position or trajectory, consistency between measured data and the expected object region, valid communication, bounded tolerances, timeout checks, and mismatch checks. If any condition is incomplete, inconsistent, or unavailable, the detection remains human-equivalent. This conservative fallback is the central safety-oriented property of inverse detection. The principle is comparable to fail-safe signal design in safety circuits: a loss of the expected condition does not preserve normal operation, but returns the system to the conservative state. For example, loss of communication with a mobile object disables its filter and causes the corresponding detections to be evaluated as human-equivalent.

Figure 1 shows the insertion of the inverse-detection module in the robot system. The module receives sensor data and known-object information, performs deterministic verification, and forwards the resulting human-equivalent detections or zone state to the safety controller.

2.2. Layered Framework

The inverse-detection framework consists of four functional layers, as shown in Figure 2:

- **Detection:** sensor measurements are collected and transformed into a com-

mon coordinate frame. All detections are initially human-equivalent.

- **Filter definition:** known system information, such as object geometry, robot pose, mobile object pose, and communication status, defines candidate filter regions.
- **Verification:** candidate filter regions are compared with measured sensor data. Filtering is permitted only when the measured data are consistent with the expected object and all verification conditions are satisfied.
- **Safety-zone evaluation:** all remaining human-equivalent detections are evaluated by the downstream safety logic.

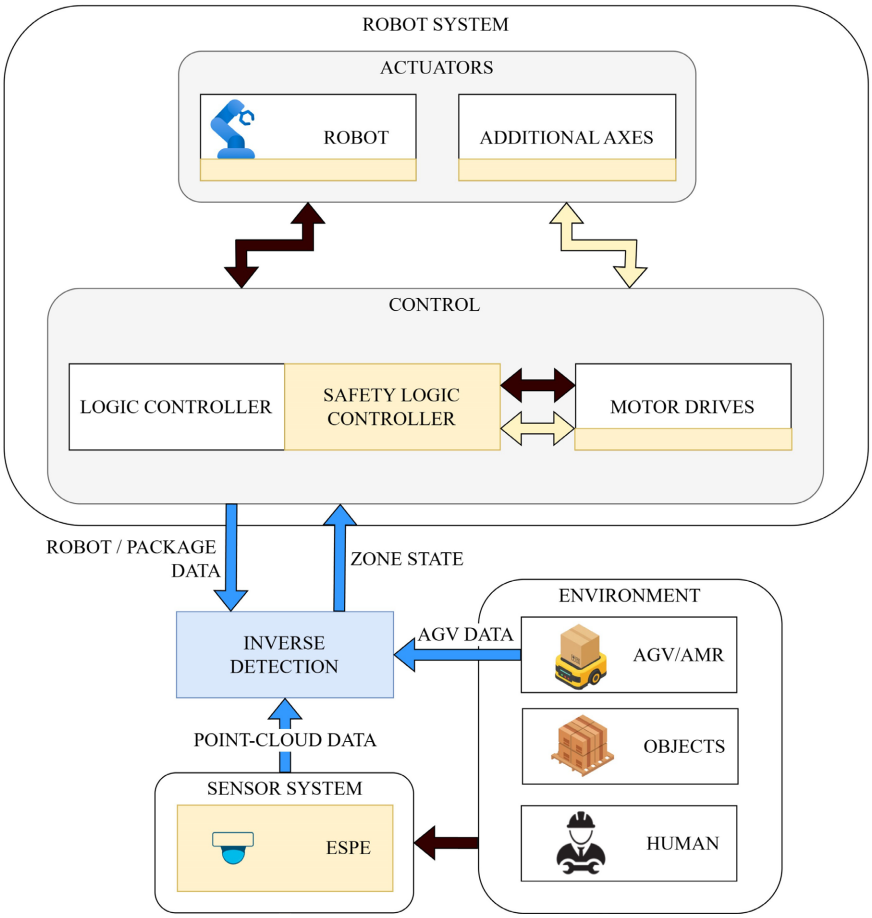

Figure 1. Insertion of the inverse-detection module in the robot system. Black arrows indicate existing functional communication, yellow arrows indicate safety-related communication, and blue arrows indicate the additional data exchange introduced by inverse detection.

In practice, the verification layers are applied within a predefined Zone of Interest (ZoI). The ZoI should at least cover the largest protective zone used by the downstream safety evaluation and may be enlarged to provide additional space for verification before detections enter the active warning or stop zones. In the proof-of-concept implementation, the ZoI was deliberately chosen larger than the active protective zones rather than being calculated as a strict minimum value.

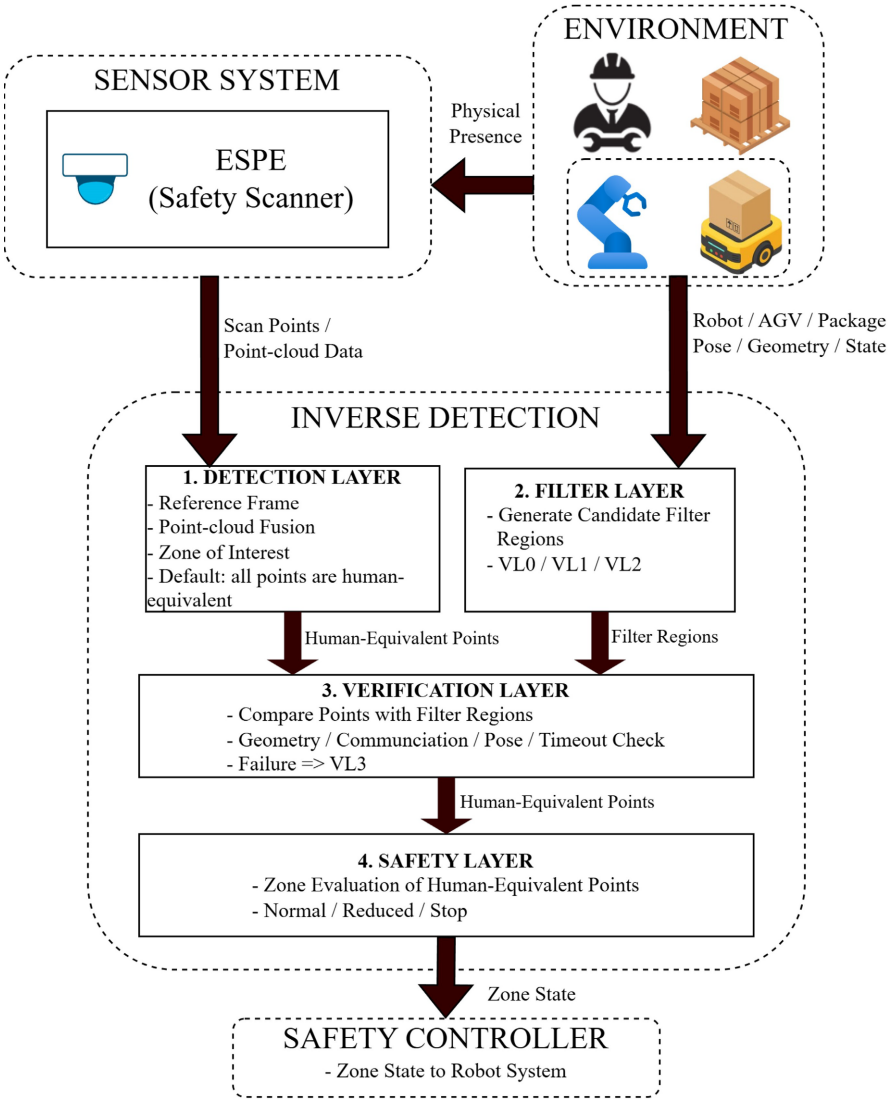

Figure 2. Dataflow and layered structure of the inverse-detection framework. Sensor measurements and known-object data enter through separate paths and are combined only at the verification stage. Only verified detections may be filtered, while unknown detections or failed verifications remain human-equivalent.

The framework is independent of the downstream zone logic. The remaining human-equivalent detections may be evaluated using fixed protective fields, ISO 13855-based safety zones, Speed and Separation Monitoring zones, or another application-specific zone-state calculation. It is also independent of the specific sensor type, as long as the sensor system provides object-position data that can be transformed into the reference frame used for verification and zone evaluation.

2.3. Verification Levels

Detected objects are handled using Verification Levels (VLs), summarized in Table 1 and illustrated in Figure 3. Each level defines the information required before detections associated with an object may be filtered.

Table 1. Verification levels used by the inverse-detection framework.

Category	Verification basis	VL	Filtering rule
Fixed structures	Known static geometry physically occupying a predefined region.	0	Detections may be filtered only inside the verified occupied volume.
Robot-manipulated objects	Known robot trajectory, object geometry, and expected object pose.	1	Detections may be filtered only while measured data match the expected object volume.
Communicating mobile objects	Communicated pose and dimensions cross-validated with sensor measurements.	2	Detections may be filtered only while communication and geometric consistency remain valid.
Human-equivalent detections	Humans, unknown objects, or objects failing verification.	3	Detections are not filtered and are forwarded to the safety-zone evaluation layer.

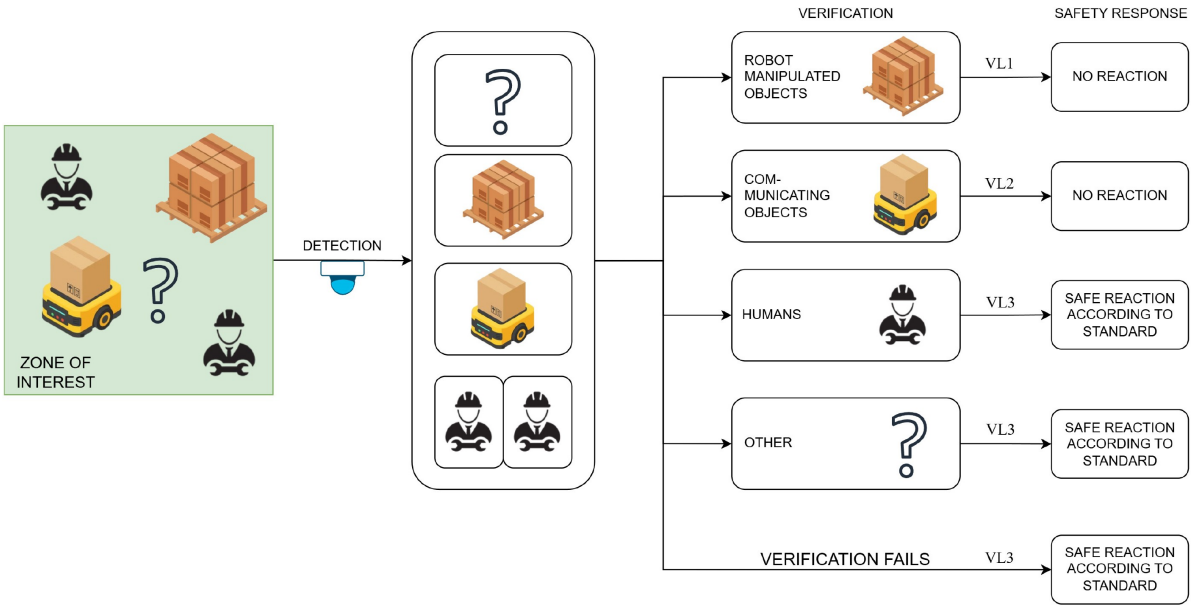

Figure 3. Simplified verification principle of inverse detection. Known non-human objects may be filtered only after deterministic verification. Humans, unknown objects, and objects failing verification are assigned VL 3 and remain part of the human-equivalent safety-zone evaluation.

VL 0 applies to fixed structures such as walls, machines, or permanently installed objects. The filter region is defined from the known geometry of the static object in the common coordinate frame. All detections inside this predefined region are removed from the safety-zone evaluation. No runtime check is performed to verify whether the region is empty or occupied. Therefore, VL 0 filters must be validated during commissioning to ensure that the static object is correctly filtered and that nearby humans or other objects are not masked. Objects assigned to VL 0 must not be removable or movable without renewed risk analysis and revalidation of the filter region.

VL 1 applies to robot-manipulated objects, for which the expected object pose is derived from the robot pose, task state, and known object dimensions. The filter region follows the expected object contour and is enlarged by a bounded geometric margin. In the proof of concept, a total margin of 100 mm was used, corresponding to 50 mm on each side of the object contour. Points inside this region are eligible for filtering, while points outside the region remain human-equivalent. If no points are detected inside the expected filter region, an empty-filter timeout is started. In the proof of concept, this timeout was set to 1 s.

VL 2 applies to communicating mobile objects, such as AGVs or AMRs. The filter region is generated from the communicated object pose and dimensions and is enlarged by the same bounded-margin principle used for VL 1. The communicated pose must be valid, timestamps must be monotonically increasing, and communication must remain active. In the proof of concept, communication validity was monitored using a watchdog signal with a timeout of 150 ms. Communication loss, watchdog timeout, stale or non-increasing timestamp information, invalid pose data, or detections outside the expected filter region invalidates the filter and returns the detections to VL 3.

VL 3 is the default level for humans, unknown objects, and all failed verifications. No filtering is applied at this level. All VL 3 detections are forwarded to the safety-zone evaluation layer and can therefore trigger the configured warning or protective-stop response. Any missing, outdated, inconsistent, or failed verification condition from VL 0-VL 2 results in VL 3 behavior.

Figure 4 illustrates this verification principle for a communicating mobile object, where the measured point cloud is compared with the communicated pose and bounded tolerance region.

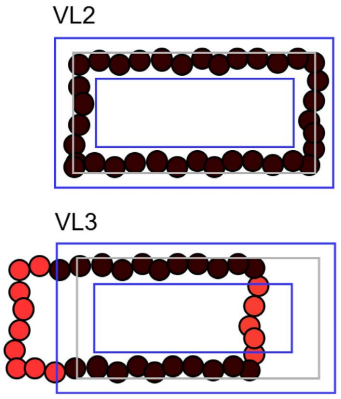

Figure 4. Verification of a communicating mobile object. The measured point cloud is compared with the communicated pose, shown as the grey box. A tolerance region, shown in blue, accounts for sensor inaccuracies and delays. Communication loss or detections outside the expected filter region trigger reclassification to VL 3.

Each candidate filter region is implemented as a deterministic state machine, as summarized in Table 2. This state formulation makes the filtering behavior explicit and prevents expected object regions from immediately removing detections

from the safety evaluation. Filtering is allowed only in the verified state. All other states preserve the corresponding detections as human-equivalent.

Table 2. Filter-state logic for candidate filter regions.

State	Meaning	Filtering allowed
Inactive	No known object is expected in the region, or the filter region is disabled.	No
Candidate	A known object is expected based on system information, but sufficient sensor confirmation has not yet been obtained.	No
Verified	The required object state is valid and the measured detections lie inside the expected filter region.	Yes
Invalid	Verification has failed due to timeout, communication loss, invalid pose data, detections outside the expected filter region, or another inconsistency.	No

A filter enters the *Candidate* state when a known object is expected and the information required to define its filter region is available, but filtering is not yet allowed. For example, a communicating mobile object may already have valid communication while it is still outside the ZoI, or a robot-manipulated package may be scheduled for transport before it is grasped by the robot. The transition from *Candidate* to *Verified* occurs when the object becomes relevant for filtering and the required conditions are satisfied, such as the mobile object entering the ZoI with valid communication, or the robot grasping the package while the expected filter region is available.

In the proof-of-concept implementation, the filter region was defined by the expected object contour with a bounded margin. The implementation did not classify the shape of the point cloud; it only checked whether detections were located inside or outside the defined filter region. A single unfiltered point inside an active safety zone was sufficient to trigger the corresponding warning or stop response.

A filter moves from *Verified* to *Invalid* when one of the required conditions is lost. For VL 1, this includes unavailable robot or task-state information, package displacement, package removal, or an empty filter region for longer than the 1 s timeout. For VL 2, this includes communication loss, watchdog timeout, stale or non-increasing timestamp information, invalid pose data, or detections outside the expected filter region. In all cases, an invalid filter no longer removes detections from the safety-zone evaluation.

2.4. Safety-Zone Integration

After verification, only detections that remain human-equivalent are forwarded to the safety-zone evaluation layer. Verified non-human objects are excluded only while their verification conditions remain valid. If verification becomes incomplete, inconsistent, or unavailable, the corresponding detections return to the human-

equivalent detection set. Depending on the application, the downstream safety logic may then maintain normal operation, reduce speed, or initiate a protective stop.

3. Proof-of-Concept Evaluation

A proof-of-concept implementation was developed to evaluate whether inverse detection can be applied in a representative industrial scenario. The goal was not certification, but a prototype-level evaluation of the conservative default: known non-human objects may be filtered only while deterministically verified, whereas humans, unknown objects, and failed verifications remain part of the safety-zone evaluation.

3.1. Experimental Setup

The testbed consisted of a Stäubli TX2-90L industrial robot mounted upside down on a horizontal linear axis, three 2D LiDAR scanners, a standard PLC for system coordination, a safety PLC for response logic, and a PC running the C++ inverse-detection algorithm. Robot pose, package dimensions, mobile object information, and zone-state data were exchanged using ADS and UDP communication.

The proof of concept used RPLIDAR A1 scanners. Their low cost enabled flexible investigation of multi-sensor coverage, placement, occlusion effects, and geometric filtering. However, they are not safety-rated and have lower measurement performance than industrial safety laser scanners. The implementation therefore demonstrates technical feasibility, not a certified safety function.

The prototype timing values were documented to characterize the proof-of-concept implementation, but they were not treated as a certified safety-response-time analysis. The RPLIDAR A1 scanners operated at approximately 10 Hz, corresponding to a scan period of about 100 ms, and transmitted their data to the PC via USB/UART. Robot and package data were transmitted to the C++ program through ADS with an expected latency below 2 ms, while AGV/AMR data were transmitted through UDP with an expected latency of approximately 2 - 5 ms. The C++ processing cycle, including point transformation, filter update, filtering, and zone-state evaluation, required approximately 90 ms in the prototype. The resulting zone state was transmitted to the PLC through ADS, after which the PLC communication update and logic response required approximately 1 ms and 10 ms, respectively. Communication loss for mobile objects was monitored separately using a watchdog timeout of 150 ms.

The three LiDAR scanners were registered to the common workspace coordinate frame using a manual calibration procedure. After mounting the scanners, several static reference objects were placed at different positions in the shared workspace, including regions visible from more than one scanner. For each scanner, a two-dimensional rigid transformation consisting of a translation and rotation was adjusted until the measured contours of the reference objects overlapped in the common workspace frame. The same workspace frame was used for the robot pose data and for the generation of the filter regions. The resulting align-

ment was verified qualitatively by comparing the registered point clouds of the reference objects and by checking that known packages and static objects were consistently located inside their expected filter regions. The remaining registration uncertainty was therefore treated as part of the filter-margin budget.

A total tolerance margin of 100 mm was added around known object geometries, corresponding to 50 mm on each side of the object contour. This margin was selected to cover scanner resolution and accuracy, registration uncertainty, coordinate-transformation error, package-position uncertainty, pose-update uncertainty, and limited object displacement during the prototype update delay. It must remain large enough for robust filtering, but small enough to avoid masking nearby human body parts or unknown objects. In the tested configuration, no complete masking of a nearby human leg was observed.

The communicating mobile object case was evaluated in two steps. First, an artificial mobile object point cloud was added to the LiDAR stream while the corresponding pose was communicated to the inverse-detection module. Second, a physical box was moved through the monitored area by the robot while its position was transmitted through UDP. This emulated the sensing and communication conditions of a communicating mobile object, without validating a complete AGV system. In the prototype, loss of UDP communication caused the corresponding filter to become invalid, so the object returned to the human-equivalent set instead of remaining filtered.

Figure 5 shows the proof-of-concept visualization, including filtered detections, remaining human-equivalent detections, and the resulting safety-zone response.

Figure 5. Visualization of the proof-of-concept implementation. Grey points represent filtered data that do not influence the zone-state evaluation. Blue and green points correspond to unfiltered data that may trigger protective actions. The human entering the red zone initiates a safety response, while verified packages in the same zone are filtered and do not affect the system response.

3.2. Evaluation Scenarios

The evaluation examined whether known non-human objects were filtered only while verified, humans remained visible to the safety-zone evaluation, and filtering was disabled under displacement, occlusion, communication loss, or geometric mismatch.

Table 3 summarizes the validation tests and repeated-trial outcomes used to evaluate the conservative behavior of the proof-of-concept implementation. A run was counted as passed when the expected filter state and safety-zone response were obtained. Detailed timing distributions of the state transitions were not logged in this proof-of-concept implementation.

Table 3. Validation tests and repeated-trial outcomes for the inverse-detection proof of concept.

Test	Expected response	Runs	Observed outcome
Human entering the monitored area	Human-equivalent points are forwarded to the safety-zone evaluation.	30	30/30 passed; no human detections were filtered.
Robot-manipulated package inside a filter region	Package points are filtered only while the measured data remain inside the expected package region.	30	30/30 passed; no unintended stops occurred during valid filtering.
Package removed from the expected region	The filter is disabled after the empty-filter timeout and the region no longer masks detections.	30	30/30 passed; the filter was revoked after the 1 s timeout; no missed reclassification occurred.
Package displaced from the expected location	Points outside the expected filter region remain human-equivalent and the filter is removed.	30	30/30 passed; the filter was revoked; no missed reclassification occurred.
Communicating mobile object inside the communicated region	The object is filtered only while communication and filter-region consistency remain valid.	30	30/30 passed; no unintended stops occurred during valid filtering.
Communication loss or invalid mobile-object data	The mobile-object filter is removed and the object becomes human-equivalent.	30	30/30 passed; the filter was revoked after watchdog timeout or invalid data; no missed reclassification occurred.
Human close to a filtered package	Human points outside the package filter region remain available for safety-zone evaluation.	30	30/30 passed in the tested configurations; the human point cloud was not completely masked.

Continued

Single-scanner occlusion and two-scanner mitigation	Occlusion is identified as a limitation; additional sensing improves visibility in the tested configuration.	Qualitative check	Single-scanner shadow regions were observed; adding a second scanner mitigated the observed occlusion in the tested setup.

3.3. Results

Across the repeated tests, humans remained human-equivalent in all runs, verified robot-manipulated packages and communicating mobile objects were filtered during valid operation, and deliberately introduced verification failures caused the expected filter revocations. No missed reclassification was observed, and no unintended stops occurred during valid filtering of verified non-human objects.

Additional tests examined whether a human standing close to a filtered package could be masked by the filter margin. In the tested configurations, the human point cloud was not completely contained within the filter region, so operator points remained available for safety-zone evaluation. This result is configuration-dependent and does not replace formal margin validation.

The combined scenario revealed an important limitation of 2D LiDAR. With a single scanner, packages at the drop-off location created shadow regions in which a human operator could be temporarily occluded. Adding a second scanner next to the package drop zone mitigated this limitation in the tested configurations.

Overall, the proof of concept showed that inverse detection can selectively filter verified non-human detections while preserving conservative behavior for humans, unknown objects, displaced objects, and communication failures. Further work is required before safety-rated deployment, including certified sensing, validated communication, timing analysis, diagnostic coverage, software verification, and formal validation of filter margins and occlusion conditions.

4. Discussion and Limitations

Inverse detection improves object selectivity by filtering only detections that can be deterministically verified as known non-human objects. All other detections remain human-equivalent. This avoids reliance on probabilistic human classification and makes the decision logic explicit, reproducible, and easier to inspect.

The approach is independent of the downstream zone logic. It does not define or reduce the required protective distance, nor does it replace the safety-zone calculation. The protective distance or zone geometry remains governed by the applicable safety concept, such as fixed protective fields, ISO 13855-based safety zones, Speed and Separation Monitoring zones, or another application-specific method. Inverse detection only determines which detections remain human-equivalent and must therefore be evaluated by the downstream safety logic. Verified non-human objects may be excluded from this evaluation while their verification conditions remain valid, whereas humans, unknown objects, and failed verifications remain active. The principle is also sensor-independent at framework level: LiDAR, radar,

or other sensors may be used if they provide sufficiently reliable position data for the detections to be verified.

The presented implementation demonstrates feasibility, but it is not a certified safety function. The RPLIDAR A1 sensors are not safety-rated, and the PC-based software and communication channels were not developed or validated according to a functional-safety lifecycle. A certified implementation would require safety-rated sensing, validated communication, timing analysis, diagnostic coverage, software verification, and formal validation of relevant failure modes.

Two technical limitations are especially important. First, 2D LiDAR creates occlusion and shadow zones behind objects, as illustrated in Figure 6. In the proof-of-concept setup, packages could occlude a human operator from a single scanner. This was mitigated by adding a second scanner near the package drop zone, but occlusion remains a fundamental limitation of 2D sensing and may require redundant sensor placement or volumetric sensing.

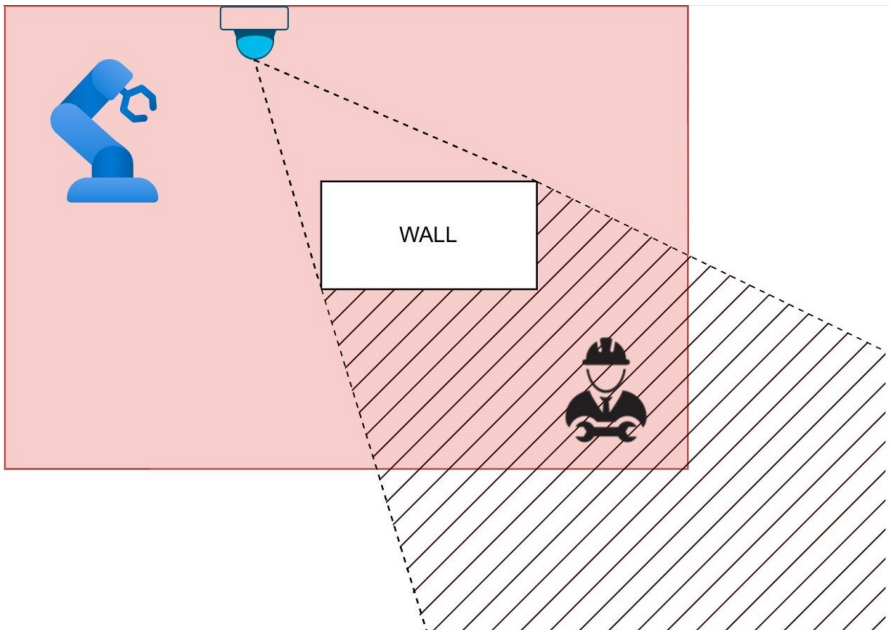

Figure 6. Illustration of shadow zones created by 2D LiDAR. Objects positioned behind other objects may be occluded from a single 2D LiDAR viewpoint. Additional sensors or volumetric sensing can mitigate this issue.

Second, the filter margin is safety-critical. A margin that is too small reduces filtering robustness, while a margin that is too large may mask nearby human body parts. The proof of concept used a 100 mm total margin, implemented as 50 mm inside and 50 mm outside the expected object contour. This value was selected from the scanner accuracy and resolution requirements of the prototype: it had to exceed the expected measurement and calibration uncertainty, while remaining small enough to preserve leg detections close to the object. No complete masking was observed under the tested configurations, including a foot placed as close as possible to the package, but this result is configuration-dependent. De-

ployment would require formal margin analysis based on sensor resolution, object dimensions, body-part dimensions, scanner placement, localization error, timing uncertainty, and worst-case occlusion geometry.

Inverse detection should therefore be interpreted as a filtering principle that supports object selectivity before safety-zone evaluation, not as a complete safety function by itself. The safety-related performance of a deployed system still depends on the complete chain: sensing, communication, computation, diagnostics, actuation, and integration with the safety controller.

5. Conclusions

This paper introduced inverse detection, a deterministic object-filtering principle for improving selectivity in zone-based industrial safety monitoring. Instead of proving that a detected object is human, the method treats all detections as human-equivalent by default and filters only detections associated with known non-human objects after deterministic verification.

The framework defines verification levels for fixed structures, robot-manipulated objects, communicating mobile systems, and human-equivalent detections. Filtering is allowed only while geometry, pose, communication status, and measured sensor data remain consistent. Humans, unknown objects, displaced objects, communication failures, and failed verifications remain human-equivalent and are forwarded to the downstream safety-zone evaluation layer.

A proof-of-concept implementation in an industrial robot cell showed that verified packages and mobile objects could be filtered from the zone-state evaluation, while humans and unknown objects remained active. The system also returned to conservative behavior when objects were displaced, filter regions became empty, or communication and geometric consistency were lost.

The implementation is not a certified safety function, but it demonstrates the feasibility of the inverse-detection principle under representative geometric and communication conditions. Future work should focus on safety-rated sensing, validated communication, timing and diagnostic analysis, formal software verification, filter-margin validation, and occlusion mitigation. Inverse detection therefore provides a practical foundation for improving object selectivity in shared industrial workspaces without relying on probabilistic human classification as the primary safety mechanism.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] International Organization for Standardization (2024) ISO 13855:2024 Safety of Machinery—Positioning of Safeguards with Respect to the Approach of the Human Body. https://www.iso.org/standard/80590.html

[2] International Organization for Standardization (2016) ISO/TS 15066:2016 Robots and Robotic Devices—Collaborative Robots.

https://www.iso.org/standard/62996.html

[3] Ali, M., Atia, M.R.A. and Fouz, M.A. (2025) Decision Enhancement of Speed and Separating Monitoring Modes for Human-Robot Collaborative Safety. *Systems and Soft Computing*, **7**, Article 200260. https://doi.org/10.1016/j.sasc.2025.200260

[4] Gómez, J., Aycard, O. and Baber, J. (2023) Efficient Detection and Tracking of Human Using 3D LiDAR Sensor. *Sensors*, **23**, Article 4720. https://doi.org/10.3390/s23104720

[5] Jia, D., Hermans, A. and Leibe, B. (2020) DR-SPAAM: A Spatial-Attention and Auto-Regressive Model for Person Detection in 2D Range Data. http://arxiv.org/abs/2004.14079

[6] Buyukakkaslar, M.T., Erturk, M.A. and Aydin, M.A. (2024) A Review on Radar-Based Human Detection Techniques. *Sensors*, **24**, Article 5709. https://doi.org/10.3390/s24175709

[7] Engel, L., Mueller, J., Rendon, E.J.F., Dorschky, E., Krauss, D., Ullmann, I., *et al.* (2025) Advanced Millimeter Wave Radar-Based Human Pose Estimation Enabled by a Deep Learning Neural Network Trained with Optical Motion Capture Ground Truth Data. *IEEE Journal of Microwaves*, **5**, 373-387. https://doi.org/10.1109/jmw.2025.3535525

[8] Zlatanski, M., Sommer, P., Zurfluh, F., Zadeh, S.G., Faraone, A. and Perera, N. (2019) Machine Perception Platform for Safe Human-Robot Collaboration. 2019 *IEEE SENSORS*, Montreal, 27-30 October 2019, 1-4. https://doi.org/10.1109/sensors43011.2019.8956547

[9] Scholz, C., Cao, H., Imrith, E., Roshandel, N., Firouzipouyaei, H., Burkiewicz, A., *et al.* (2025) Sensor-Enabled Safety Systems for Human-Robot Collaboration: A Review. *IEEE Sensors Journal*, **25**, 65-88. https://doi.org/10.1109/jsen.2024.3496905

[10] Oleinikov, A., Soltan, S., Balgabekova, Z., Bemporad, A. and Rubagotti, M. (2024) Scenario-Based Model Predictive Control with Probabilistic Human Predictions for Human-Robot Coexistence. *Control Engineering Practice*, **142**, Article 105769. https://doi.org/10.1016/j.conengprac.2023.105769

[11] Kamezaki, M., Wada, T. and Sugano, S. (2024) Dynamic Collaborative Workspace Based on Human Interference Estimation for Safe and Productive Human-Robot Collaboration. *IEEE Robotics and Automation Letters*, **9**, 6568-6575. https://doi.org/10.1109/lra.2024.3405352

[12] Awele, O., Odili, J., Olukoya, D., C Onwuegbuchi, O., Altawati, M. and Agbeso, D.O. (2026) Safe and Explainable Artificial Intelligence for Safety-Critical Robotic Systems. *International Journal of Science and Research Archive*, **18**, 167-176. https://doi.org/10.30574/ijsra.2026.18.1.0034

[13] Arents, J., Abolins, V., Judvaitis, J., Vismanis, O., Oraby, A. and Ozols, K. (2021) Human-Robot Collaboration Trends and Safety Aspects: A Systematic Review. *Journal of Sensor and Actuator Networks*, **10**, Article 48. https://doi.org/10.3390/jsan10030048

[14] International Organization for Standardization (2023) ISO 13849-1: 2023 Safety of Machinery—Safety-Related Parts of Control Systems—Part 1: General Principles for Design. https://www.iso.org/standard/73481.html

[15] Perez-Cerrolaza, J., Abella, J., Borg, M., Donzella, C., Cerquides, J., Cazorla, F.J., *et al.* (2024) Artificial Intelligence for Safety-Critical Systems in Industrial and Transportation Domains: A Survey. *ACM Computing Surveys*, **56**, 1-40. https://doi.org/10.1145/3626314

[16] International Organization for Standardization (2024) ISO/IEC TR 5469: 2024 Arti-

ficial Intelligence—Functional Safety and AI Systems.
https://www.iso.org/standard/81283.html

[17] Magrini, E., Ferraguti, F., Ronga, A.J., Pini, F., De Luca, A. and Leali, F. (2020) Human-Robot Coexistence and Interaction in Open Industrial Cells. *Robotics and Computer-Integrated Manufacturing*, **61**, Article 101846.
https://doi.org/10.1016/j.rcim.2019.101846

[18] Malm, T., Salmi, T., Marstio, I. and Montonen, J. (2019) Dynamic Safety System for Collaboration of Operators and Industrial Robots. *Open Engineering*, **9**, 61-71.
https://doi.org/10.1515/eng-2019-0011

[19] Alenjareghi, M.J., Keivanpour, S., Chinniah, Y.A., Jocelyn, S. and Oulmane, A. (2024) Safe Human-Robot Collaboration: A Systematic Review of Risk Assessment Methods with AI Integration and Standardization Considerations. *The International Journal of Advanced Manufacturing Technology*, **133**, 4077-4110.
https://doi.org/10.1007/s00170-024-13948-3

[20] Bermuth, D., Poeppel, A. and Reif, W. (2025) Tutabo-1: Towards Real-Time Capable AI-Based Safety Systems for Human-Robot Collaboration. 2025 *IEEE International Conference on Advanced Robotics* (*ICAR*), San Juan, 2-5 December 2025, 632-639.
https://doi.org/10.1109/icar65334.2025.11338695