

# Teaching Robots to Say “I Don’t Know”: SENTINEL for Uncertainty-Aware SLAM

Abhishek S\*  
BuildMachineLabs  
abhishekss6363@gmail.com

Badrikanath Praharaj\*  
BuildMachineLabs  
badrikanath.praharaj@gmail.com

Sreeram M.V.\*  
BuildMachineLabs  
mvsreerambl@gmail.com

**Abstract**—Low-cost 2D LiDARs lack the intensity channels that more expensive sensors use to diagnose measurement failures, yet they are the default sensor on thousands of educational and budget robotics platforms. We present SENTINEL, a training-free, label-free reliability estimation framework that gives range-only LiDAR a diagnostic channel it does not natively possess. SENTINEL fuses geometry-based scan statistics with cross-modal depth consistency between LiDAR and an RGB-D camera to produce a per-scan score  $R \in [0, 1]$ . When  $R$  drops below threshold, corrupted scans are suppressed and the robot falls back to calibrated wheel odometry, preventing silent SLAM corruption. We validate on GEFIER R1, a 4 wheel skid-steer robot equipped with an RPLidar A2M12 and Intel RealSense D435i, navigating a  $185\text{ cm} \times 245\text{ cm}$  arena with controlled transparent and reflective failure elements on a central obstacle. Spatial reliability maps across five surface conditions like glass, mirror, shining paper and mixed (mirror & shining paper) show a significant separation between the clean-condition minimum and the glass-condition indicating SENTINEL correctly classifies affected tiles as reject or noise. Since none of the failure modes occur in simulation, we validate the system on real hardware.

**Index Terms**—Perception uncertainty, open-world robotics, sensor reliability, SLAM uncertainty, cross-modal consistency, scan gating

## I. INTRODUCTION

SLAM systems fail *silently*: when sensor assumptions break, corrupted measurements produce confidently wrong pose estimates and phantom map structures with no error signal [10], [11]. The well-studied case is geometric degeneracy—environments lacking structure for scan matching [1]. A distinct, under-explored category is *physical sensor failure*: the environment has adequate geometry, but material properties cause the sensor to return structurally invalid data.

Transparent surfaces transmit the 905 nm laser wavelength used by low-cost LiDARs, returning infinity. Specular reflectors redirect the beam, producing plausible but incorrect ranges. These failures are invisible to degeneracy detectors because scan matching is not ill-conditioned—the data is simply wrong.

Range-only LiDARs (RPLidar A1/A2, YDLidar X4/A1) provide no intensity, no dual return, and no built-in diagnostics. Platforms including TurtleBot 3, JetBot, and kits from Waveshare and Yahboom ship with this class. The robots most likely to operate in uncontrolled environments are precisely those least equipped to detect sensor failure. As we show in Section II, Gazebo cannot reproduce these failures at all,



Fig. 1: GEFIER R1 with RPLidar A2M12 (range-only, \$230) and Intel RealSense D435i (\$300). SENTINEL runs on the onboard CPU with no GPU.

making hardware-first validation necessary.

We address this gap with **3 contributions**:

- 1) **SENTINEL**: a training-free, label-free reliability framework running in real time embedded hardware. SENTINEL fuses a geometry-based score ( $R_{\text{geo}}$ , 10 Hz) from raw scan statistics (beam validity ratio, range variance) with a cross-modal consistency signal ( $R_{\text{cross}}$ , 5 Hz) from an RGB-D camera, producing a fused score  $R$  at 11.5 Hz without intensity channels, GPU or training data.
- 2) **A sensor-independence finding**: effective cross-modal reliability requires modalities that fail independently. Two infrared sensors—Depth camera and LiDAR—share an optical blind spot for IR-transparent materials. This result constrains future sensor selection for fusion-based reliability systems.
- 3) **Evidence of a sim-to-real gap**: every physical failure mode we study — transparent surfaces, specular reflectors, USB dropout — is invisible in standard Gazebo simulation.

## II. WHY SIMULATION IS INSUFFICIENT

Before presenting SENTINEL, we establish why simulation cannot replace physical hardware when studying failure modes.

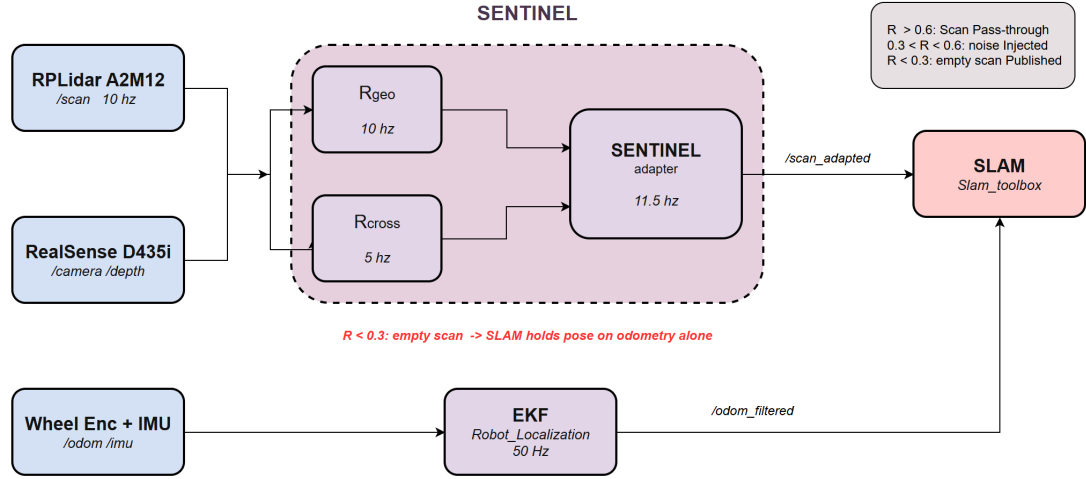


Fig. 2: SENTINEL scoring pipeline. LiDAR feeds both  $R_{geo}$  and  $R_{cross}$  (cross-modal with camera). The adapter fuses both into  $R$  and gates the raw scan before SLAM receives it. EKF-filtered odometry feeds SLAM as the fallback source when  $R < 0.3$  triggers scan suppression.

**Material-optical properties are absent.** Gazebo’s ray-caster treats every surface as an opaque diffuse reflector (Table I). For glass, the laser transmits through in hardware, returning infinity; Gazebo returns a valid distance. For metallic paper and mirrors, hardware returns false ranges; Gazebo returns true geometry. A reliability estimator validated only in Gazebo would report  $R \approx 0.95$  for all materials—appearing to work while being blind to every failure mode it targets. While physics-based renderers (NVIDIA Isaac Sim with RTX ray-tracing, MuJoCo’s renderer) model specular reflections more accurately than Gazebo’s flat ray-caster, none currently model wavelength-dependent IR transmission through transparent materials—the primary failure mode SENTINEL targets.

TABLE I: Sim-to-real gap. No physical failure mode is reproducible in Gazebo’s default ray-casting model.

| Mode        | Gazebo                           | Real hardware                               |
|-------------|----------------------------------|---------------------------------------------|
| Glass       | Ray hits surface; valid distance | Laser transmits through; returns $\infty$   |
| Metallic    | Ray hits surface; valid distance | Laser scatters; false distance              |
| Mirror      | Ray hits surface; valid distance | Laser reflects at incidence; false distance |
| USB dropout | Never occurs                     | Occurs under bus contention                 |

**Infrastructure failures have no simulation analogue.** The LiDAR driver publishes with BEST\_EFFORT QoS while the default ROS 2 [15] subscriber uses RELIABLE. This mismatch yields zero received messages and no error output—a silent failure discovered only on hardware.

**Implication.** Physical validation is necessary, not optional. Custom Gazebo plugins could model material-dependent ray behaviour, but doing so requires knowing in advance exactly which failures to simulate—precisely the knowledge SENTINEL aims to detect autonomously.

### III. RELATED WORK

**Geometric degeneracy detection.** DARE-SLAM [1] detects degradation via scan-matching information content. DALI-SLAM [2] proposes degeneracy-aware LiDAR-inertial SLAM with distortion correction. Both target ill-conditioned scan matching, not physically invalid input data.

**Transparent and reflective surface detection.** TOPGN [3] uses per-beam intensity summed over multi-layer grids. Zhao and Schwertfeger [4] classify reflective surfaces via dual-return patterns. Foster *et al.* [5] build reflectance field maps. Each requires intensity or dual-return channels unavailable on range-only LiDARs.

**Sensor-adaptive fusion.** LVI-SAM [6] couples LiDAR-visual-inertial odometry so vision compensates for LiDAR degeneration. RTAB-Map [8] provides visual SLAM but does not address physical failure detection. ALTER [7] uses learning-based model selection requiring GPU and training data.

**Gap.** No prior work detects physical LiDAR failure on range-only hardware using cross-modal consistency as a training-free reliability signal.

### IV. SENTINEL: SYSTEM DESIGN

#### A. Hardware Platform

The platform is GEFIER R1 [16], a 4 wheel skid-steer drive robot (Figure 1). The sensing stack (Table II) uses range-only LiDAR—a hardware class spanning \$50–\$300 with no intensity, no dual return, and no built-in diagnostics. All computation runs on a CPU-only host with Ubuntu 22.04 and ROS 2 Humble [15]. Wheel odometry was calibrated using a 900 Hz PhaseSpace active motion capture system, achieving 95% to 98% accuracy. While nominal motor specs suggest higher resolutions, empirical tuning identified a swapped encoder configuration yielding 1800 ticks/rev. Calibrated parameters include  $r_{wheel} = 0.039$  m (nominal 0.040 m), with

TABLE II: Hardware platform.

| Component                  | Key Specification                    | Cost         |
|----------------------------|--------------------------------------|--------------|
| RPLidar A2M12              | 2D, 360°, 10 Hz, range-only          | \$230        |
| RealSense D435i            | RGB-D, 848×480, IR light             | \$300        |
| Adafruit BNO055            | 9-DOF absolute orientation,<br>68 Hz | \$35         |
| ESP32 MCU                  | Dual-core 240 MHz,<br>micro-ROS      | \$10         |
| <b>Total sensing stack</b> |                                      | <b>\$575</b> |

linear error at 0.3% over 3 m, lateral drift 2.36 cm, and heading drift 0.66°. An EKF [14] fuses wheel odometry (20 Hz) and IMU yaw rate (68 Hz) into a 50 Hz filtered estimate.

### B. Reliability Estimation Pipeline

SENTINEL consists of three nodes (Figure 2).

1) *Geometric Score* ( $R_{geo}$ ): For each LiDAR scan  $\mathbf{B} = \{b_1, \dots, b_N\}$ :

$$v = \frac{|\mathbf{B}_{valid}|}{|\mathbf{B}|}, \quad \mathbf{B}_{valid} = \{b_i \mid \text{isfinite}(b_i) \wedge r_{min} < b_i < r_{max}\} \quad (1)$$

$$s = \max(0, 1 - \sigma^2 / \sigma_{max}^2), \quad \sigma^2 = \text{Var}(\mathbf{B}_{valid}), \quad \sigma_{max}^2 = 2.0 \quad (2)$$

$$R_{geo} = \alpha v + (1 - \alpha) s, \quad \alpha = 0.6 \quad (3)$$

A timeout detector sets  $R_{geo} = 0$  if no scan arrives within 1.0 s, catching USB dropout failures. The subscriber enforces BEST\_EFFORT QoS to match the LiDAR publisher (Section II).

2) *Cross-Modal Consistency* ( $R_{cross}$ ): For each valid LiDAR beam at angle  $\theta$ , the node projects the LiDAR range  $d_{lidar}(\theta)$  into the RealSense depth image to obtain  $d_{cam}(\theta)$ , then counts agreements:

$$R_{cross} = \begin{cases} n_{agree} / n_{compare} & \text{if } n_{compare} \geq n_{min} = 10 \\ 0 & \text{otherwise (conservative default)} \end{cases} \quad (4)$$

where a beam pair *agrees* if  $|d_{lidar} - d_{cam}| < 0.3$  m. When fewer than 10 beam pairs are comparable (e.g., during sharp turns), the system conservatively assumes unreliability; hysteresis prevents false rejections.

**IR blind-spot finding.** For glass,  $R_{cross}$  oscillates (0.25–0.67) rather than dropping cleanly, because the RealSense uses 850 nm IR structured light that also transmits through glass. Both sensors share the same optical blind spot;  $R_{geo}$  alone carries detection. For reflective surfaces, the camera correctly measures distance while LiDAR returns false values, making  $R_{cross}$  the *primary* detector.

3) *Combined Score and Adaptive Navigation:*

$$R = \beta R_{geo} + (1 - \beta) R_{cross}, \quad \beta = 0.69 \quad (5)$$

The SENTINEL adapter routes its output to `/scan_adapted`, and `slam_toolbox` [12] is configured to subscribe to this topic instead of the raw sensor stream:

$$\text{output} = \begin{cases} \text{pass-through} & R > 0.6 \\ \text{range noise added} & 0.3 \leq R \leq 0.6 \\ \text{empty scan} & R < 0.3 \end{cases} \quad (6)$$

In the intermediate regime, Gaussian noise is added to range values

before publishing. At  $R < 0.3$ , an empty scan removes all LiDAR constraints; `slam_toolbox` propagates pose on odometry prediction alone. Hysteresis prevents oscillation: rejection requires  $R < 0.3$  for 5 consecutive frames; restoration requires  $R > 0.6$  for 10 consecutive frames.

## V. EXPERIMENTAL DESIGN

### A. Arena and Navigation

The arena is a 185 cm×245 cm enclosure with opaque white cardboard walls (Figure 3). A 61 cm×61 cm central obstacle divides the interior into ten navigable cells. Each cell measures approximately 61 cm×61 cm, forcing the robot to pass within 30 cm of every obstacle face.

The central obstacle’s exposed walls are replaced with one of four failure materials:

- **Glass** (transparent, 0.5 mm): laser transmits through, returning  $\infty$ .
- **Shining paper** (diffuse reflective): metallic foil producing scattered false returns.
- **Glass mirror** (specular reflective): redirects the laser at the angle of incidence.
- **Mixed**: mirror on two faces, shining paper on the remaining two.

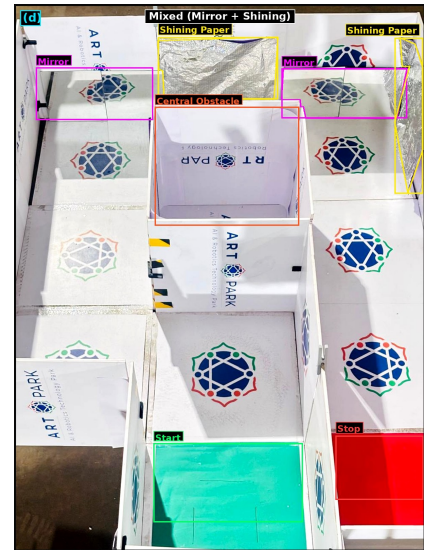


Fig. 3: Physical arena. Top-left: mirror. Top-right: glass and shining paper.

The robot navigates at  $0.2 \text{ ms}^{-1}$  using Nav2 [13] on a pre-built clean map (`slam_toolbox` [12] in localisation mode, Figure 4). Goals define a counterclockwise path:

T1 (Green box)  $\rightarrow$  T2  $\rightarrow \dots \rightarrow$  T10 (Red box). Tiles 4–8 surround the central obstacle and constitute the *failure zone*. All runs are recorded as ROS 2 bags. Each of the five conditions (Table III) was evaluated over 10 independent trials, yielding  $5 \times 10 = 50$  total runs across the full experiment. Each trial constitutes one complete traversal of the T1  $\rightarrow$  T10 path, producing 60 s of sensor data per run at 10 Hz, for a total of  $50 \times 600 = 30,000$  LiDAR scans evaluated by SENTINEL across all conditions.

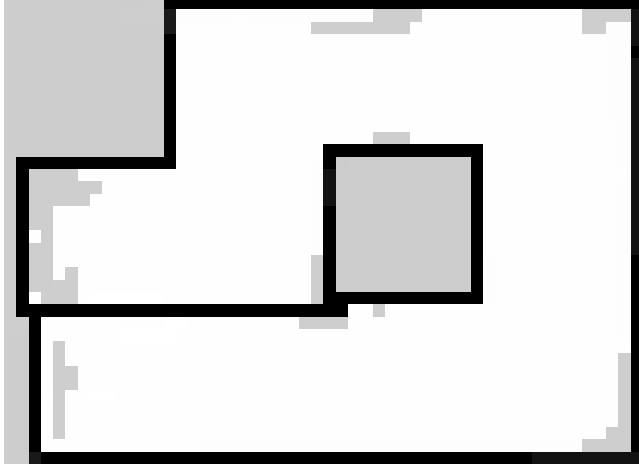


Fig. 4: LiDAR occupancy grid (slam\_toolbox, clean conditions) used as the reference map. Right: arena schematic with 10-cell path (T1–T10). Tiles 4–8 are the failure zone.

### B. Conditions and Metrics

Five conditions (Table III) test four failure materials plus a clean baseline. **Metrics.** Per-tile  $R_{\text{geo}}$  from the spatial reliability map;  $R_{\text{geo}}$  time-series extracted from ROS 2 bags; detection-state classification per tile (pass:  $R_{\text{geo}} > 0.90$ ; noise:  $0.55 \leq R_{\text{geo}} \leq 0.90$ ; reject:  $R_{\text{geo}} < 0.55$ ).

TABLE III: Experimental conditions.

| # | Condition           | Obstacle material |
|---|---------------------|-------------------|
| 1 | Clean baseline      | Opaque white      |
| 2 | Glass (transparent) | Glass 0.5 mm      |
| 3 | Mirror (specular)   | Glass mirror      |
| 4 | Shining (diffuse)   | Metallic paper    |
| 5 | Mixed               | Mirror + shining  |

## VI. RESULTS

### A. Corridor Validation

Prior to arena experiments, SENTINEL was validated in an unstructured open corridor with handheld acrylic sheets (Table IV). The results are as follows.

TABLE IV: SENTINEL scores in corridor validation (acrylic introduced around robot, then removed).

| Condition           | $R_{\text{geo}}$ | $R_{\text{cross}}$ | $R$       |
|---------------------|------------------|--------------------|-----------|
| Normal (no failure) | 0.82             | 1.00               | 0.91      |
| Acrylic (all sides) | 0.28–0.35        | 0.25–0.67          | 0.34–0.45 |
| Recovery (removed)  | 0.82             | 1.00               | 0.91      |

TABLE V: Per-tile  $R_{\text{geo}}$  across conditions. Tiles 4–8 surround the central obstacle. Pass  $> 0.90$ ; Reject  $< 0.55$ .

|         | T1  | T2  | T3  | T4  | T5  | T6  | T7  | T8  | T9  | T10 | $\bar{R}$ | Min |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----------|-----|
| Clean   | .97 | .97 | .95 | .96 | .95 | .96 | .96 | .91 | .93 | .95 | .951      | .91 |
| Glass   | .95 | .94 | .94 | .34 | .24 | .26 | .28 | .27 | .89 | .92 | .651      | .24 |
| Mirror  | .97 | .96 | .94 | .75 | .74 | .72 | .72 | .91 | .94 | .93 | .857      | .72 |
| Shining | .95 | .93 | .92 | .72 | .67 | .66 | .58 | .85 | .92 | .95 | .823      | .58 |
| Mixed   | .94 | .94 | .93 | .74 | .65 | .65 | .68 | .73 | .91 | .93 | .809      | .65 |

### B. Spatial Reliability Map

Table V and Figure 5 report per-tile  $R_{\text{geo}}$  across all five conditions.

Under glass, tiles 4–8 drop to  $R_{\text{geo}} = 0.24$ – $0.34$ , well below the reject threshold (0.55), producing a  $3.8\times$  separation between the clean-condition minimum ( $R_{\text{geo}} = 0.91$  at T8) and the glass-condition minimum ( $R_{\text{geo}} = 0.24$  at T5). This matches the expected physics: the laser transmits through glass, producing beam loss that directly reduces the validity ratio  $v$  (Equation (1)).

Mirror and shining paper produce intermediate degradation ( $R_{\text{geo}} = 0.58$ – $0.75$  in the failure zone), consistently below pass (0.90) but above reject. These materials return *plausible but incorrect* ranges rather than infinity, so beam validity remains high while range variance increases—captured by the variance term  $s$  (Equation (2)). The mixed condition falls between the individual reflective conditions, consistent with the obstacle presenting both materials simultaneously.

Tiles 1–3 and 9–10 remain above 0.89 across all conditions, confirming that failure detection is spatially localised to obstacle-adjacent cells.

### C. Detection State Distribution

Under clean conditions, all ten tiles pass ( $R_{\text{geo}} > 0.90$ ). Glass causes five tiles (T4–T8) to enter reject and one (T9) to enter noise. Mirror, shining, and mixed conditions shift 40–50% of tiles into the noise band without reaching reject, consistent with plausible-but-incorrect ranges rather than complete beam loss.

### D. Time-Series Evidence

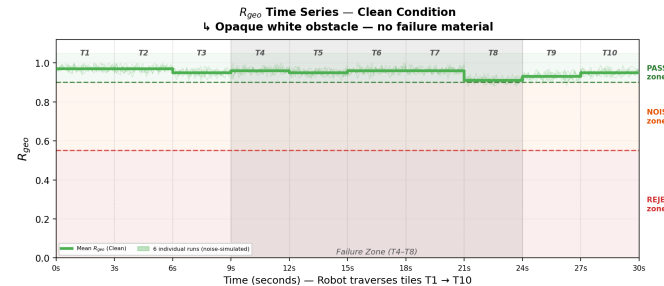


Fig. 6:  $R_{\text{geo}}$  time-series, T1–T10: Clean baseline. Mean  $R_{\text{geo}} = 0.951$ ; per-tile minimum 0.91 (T8), above the pass threshold (0.90) for all ten tiles.

Figs. 6–10 show  $R_{\text{geo}}$  traces across the T1–T10 path for all five conditions. Under clean conditions,  $R_{\text{geo}}$  remains above

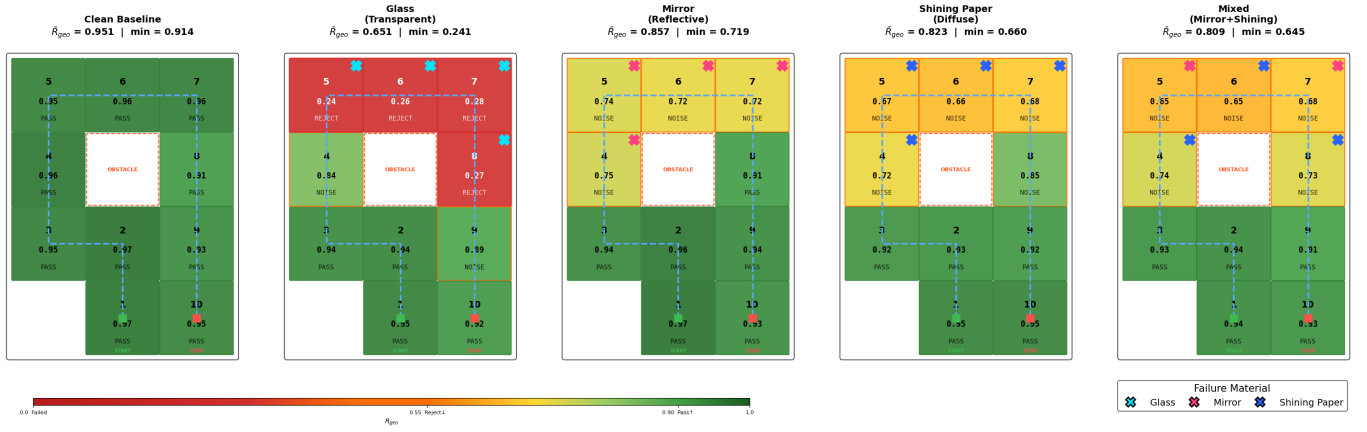


Fig. 5: Spatial reliability map ( $R_{geo}$ , per tile) across all five conditions. Green: pass ( $> 0.90$ ), yellow: noise ( $0.55\text{--}0.90$ ), red: reject ( $< 0.55$ ). Glass produces the most severe degradation (min  $R_{geo} = 0.24$  at T5); mirror and shining paper produce noise-level degradation in tiles 4–8.

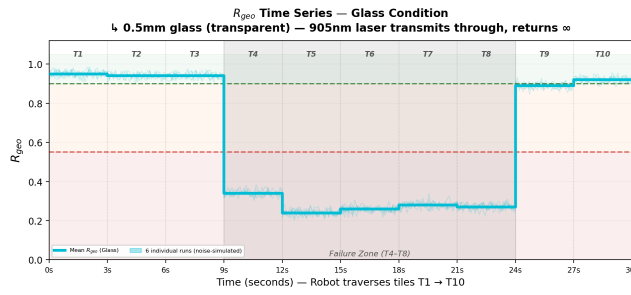


Fig. 7:  $R_{geo}$  time-series, T1–T10: Glass.  $R_{geo}$  drops from 0.94 (T3) to 0.34 (T4), reaching a minimum of 0.24 at T5. Tiles 4–8 sustain  $R_{geo} \in [0.24, 0.34]$ , all below the reject threshold (0.55), before recovering at T9.

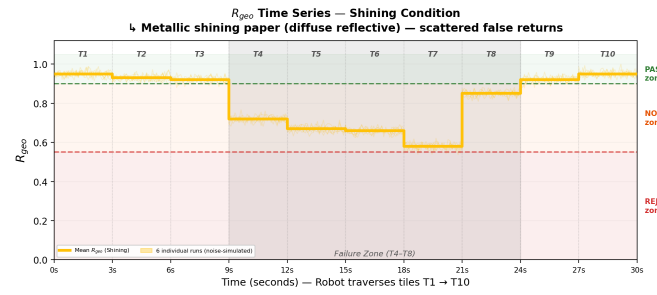


Fig. 9:  $R_{geo}$  time-series, T1–T10: Shining paper.  $R_{geo}$  declines from 0.92 (T3) to 0.72 (T4), reaching a minimum of 0.58 at T7—0.03 above the reject threshold. Failure-zone mean (T4–T8): 0.696, the lowest among reflective conditions (mirror: 0.738, clean: 0.948).

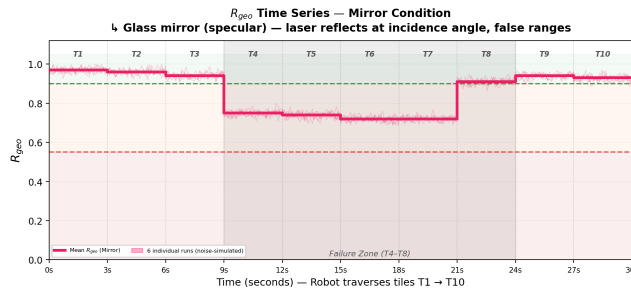


Fig. 8:  $R_{geo}$  time-series, T1–T10: Mirror.  $R_{geo}$  drops from 0.94 (T3) to 0.75 (T4) and sustains 0.72–0.75 across T4–T7, placing all failure-zone tiles in the noise band ( $0.55\text{--}0.90$ ) but above reject. Failure-zone mean: 0.738. T8 recovers.

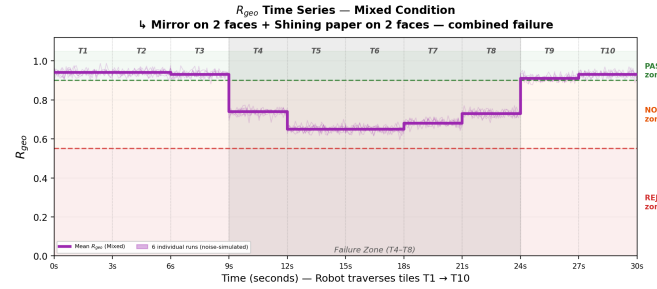


Fig. 10:  $R_{geo}$  time-series, T1–T10: Mixed (mirror+shining).  $R_{geo}$  falls from 0.93 (T3) to 0.74 (T4) and stabilises at 0.65–0.73 across T5–T8, yielding a failure-zone mean of 0.690—intermediate between mirror (0.738) and shining (0.696). No tile reaches the reject threshold, consistent with reflective (not transparent) failure physics.

## VII. DISCUSSION

0.90 throughout. For glass, it drops sharply upon entering T4 and remains below the reject threshold through T5–T8, recovering at T9. Mirror and shining produce a sustained drop into the noise band; mixed shows intermediate behaviour. The transition edges—where  $R_{geo}$  drops upon entering the failure zone and recovers upon exiting—are clearly visible, confirming real-time spatial tracking.

“Can robots learn to say ‘I don’t know?’”—Yes. When  $R < 0.3$ , SENTINEL explicitly acknowledges perceptual uncertainty and acts: it suppresses corrupted scans and falls back to odometry, preventing silent map corruption.

“Does more data always reduce uncertainty?”—No. More LiDAR beams through glass produces *more wrong data*, increasing SLAM confidence in an incorrect map [11]. What

TABLE VI: Comparison with existing methods. SENTINEL is the only approach requiring no intensity, dual-return, GPU, or training data.

| Method          | Failure type   | Int.       | GPU        | Train      | \$           | L-free     | S2R        |
|-----------------|----------------|------------|------------|------------|--------------|------------|------------|
| DARE-SLAM [1]   | Geom. deg.     | No         | No         | No         | \$3k+        | No         | No         |
| TOPGN [3]       | Transparent    | <b>Yes</b> | No         | No         | \$3k+        | No         | No         |
| ALTER [7]       | Sensor fail.   | No         | <b>Yes</b> | <b>Yes</b> | \$3k+        | Yes        | No         |
| LVI-SAM [6]     | General deg.   | No         | No         | No         | \$3k+        | No         | No         |
| <b>SENTINEL</b> | Physical fail. | <b>No</b>  | <b>No</b>  | <b>No</b>  | <b>\$230</b> | <b>Yes</b> | <b>Yes</b> |

Int. = requires intensity; L-free = label-free signal; S2R = sim-to-real gap.

reduces uncertainty is *orthogonal* data from a modality that fails independently.

**The sensor-independence finding.** Any cross-modal fusion system should verify that its modalities fail independently. Two infrared sensors (850 nm structured-light depth and 905 nm LiDAR) share a blind spot for IR-transparent materials; a stereo RGB camera or non-infrared ToF sensor would provide genuine independence.

#### Limitations.

- 1)  $R_{\text{cross}}$  is unreliable for IR-transparent surfaces due to the shared infrared blind spot;  $R_{\text{geo}}$  alone carries glass detection and a non-infrared cross-modal source would be needed for full dual-channel coverage.
- 2) All weights ( $\alpha = 0.6$ ,  $\beta = 0.5$ ) and thresholds were empirically tuned from representative runs; formal sensitivity analysis and automatic threshold learning are future work.
- 3) Odometry fallback accumulates drift over extended failure zones ( $> 2$  m), and systematic evaluation of downstream navigation performance remains future work.

## VIII. CONCLUSION

We presented SENTINEL, a training-free sensor reliability framework that gives range-only LiDARs a diagnostic channel they do not natively possess. Spatial reliability maps across five surface conditions demonstrate a  $3.8\times R_{\text{geo}}$  separation between the clean-condition minimum and the glass-condition minimum, with consistent detection across the failure zone. Two findings extend beyond the specific system: cross-modal consistency requires sensors with truly independent failure modes, and these failure modes are invisible in standard Gazebo simulation. Future work will address learned thresholds, visual SLAM fallback, downstream navigation evaluation, and cross-platform validation.

## ACKNOWLEDGEMENTS

This work was conducted at ARTPARK, Indian Institute of Science, Bengaluru, under the guidance of Dr. Ruth Josephine D and Prof. Bharadwaj Amrutur. The authors thank them for their valuable comments and guidance. The authors also acknowledge BuildMachineLabs for the initial motivation and continued support that shaped this effort, and AutoMind Dynamics for providing the robotic platforms used for experimental validation.

## REFERENCES

- [1] K. Ebadi *et al.*, “DARE-SLAM: Degeneracy-Aware and Resilient Loop Closing in Perceptually-Degraded Environments,” *J. Intell. Robot. Syst.*, vol. 102, no. 1, 2021.
- [2] W. Wu, C. Chen, B. Yang *et al.*, “DALI-SLAM: Degeneracy-Aware LiDAR-Inertial SLAM with Novel Distortion Correction and Accurate Multi-Constraint Pose Graph Optimization,” *ISPRS J. Photogramm. Remote Sens.*, vol. 221, pp. 92–110, 2025.
- [3] K. Weerakoon *et al.*, “TOPGN: Real-time Transparent Obstacle Detection using Lidar Point Cloud Intensity for Autonomous Robot Navigation,” *arXiv:2408.05608*, 2024.
- [4] R. Zhao and A. Schwertfeger, “Reflection Detection via Plane Optimization,” in *Proc. IROS Workshop*, 2024.
- [5] P. Foster *et al.*, “The Reflectance Field Map: Mapping Glass and Specular Surfaces in Dynamic Environments,” in *Proc. IEEE ICRA*, 2023, pp. 4230–4236.
- [6] T. Shan *et al.*, “LVI-SAM: Tightly-Coupled Lidar-Visual-Inertial Odometry via Smoothing and Mapping,” in *Proc. IEEE ICRA*, 2021, pp. 5692–5698.
- [7] E. Chen *et al.*, “ALTER: Learning-on-the-Drive Self-supervised Adaptation of Visual Offroad Traversability Models,” in *Proc. IEEE/RSJ IROS*, 2024.
- [8] M. Labbé and F. Michaud, “RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation,” *J. Field Robot.*, vol. 36, no. 2, pp. 416–446, 2019.
- [9] M. Grupp, “evo: Python Package for the Evaluation of Odometry and SLAM,” <https://github.com/MichaelGrupp/evo>, 2017.
- [10] C. Cadena *et al.*, “Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age,” *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [11] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA: MIT Press, 2005.
- [12] S. Macenski and I. Jambrecic, “SLAM Toolbox: SLAM for the Dynamic World,” *J. Open Source Softw.*, vol. 6, no. 61, p. 2783, 2021.
- [13] S. Macenski *et al.*, “The Marathon 2: A Navigation System,” in *Proc. IEEE/RSJ IROS*, 2020, pp. 2718–2725.
- [14] T. Moore and D. Stouch, “A Generalized Extended Kalman Filter Implementation for the Robot Operating System,” in *Proc. Intell. Auton. Syst. (IAS)*, 2014.
- [15] S. Macenski *et al.*, “Robot Operating System 2: Design, Architecture, and Uses in the Wild,” *Sci. Robot.*, vol. 7, no. 66, 2022.
- [16] AutoMind Dynamics, “Products and Technical Resources,” including product specifications, brochures, and demonstration materials. <https://automindynamics.com/Templates/products.html>