Check for updates

RESEARCH ARTICLE

Secure3D-CV: A scientific framework for integrity assessment of 3D data with an external Python library

[version 1; peer review: awaiting peer review]

Karol Kozak[1,2]

[1]University of Lower Silesia, Wrocław, Lower Silesian Voivodeship, Poland
[2]Q4Pro Sp.z.o.o, Bydgoszcz, Poland

v1 **First published:** 09 Jan 2026, **6**:19
https://doi.org/10.12688/openreseurope.22181.1
Latest published: 09 Jan 2026, **6**:19
https://doi.org/10.12688/openreseurope.22181.1

Abstract

Background

Three-dimensional (3D) perception is widely used in robotics, autonomous vehicles and medicine. These systems often operate autonomously and are vulnerable to data tampering or spoofing attacks. Existing work on 2D adversarial detection does not generalise well to the high-dimensional and structured nature of 3D point clouds.

Methods

We propose Secure3D-CV, a method that builds on OpenCV by adding a library for outlier detection and threat assessment in 3D scenes. The method transforms each scene (S) into a feature vector ($x = \varphi(S)$) via geometric, spectral and volumetric descriptors. A statistical anomaly score (A (S) = | x — { x ¯ } | 2) measures the deviation of (x) from a benign distribution ({ x ¯ }). A supervised or one-class classifier ($f_\theta(x)$) estimates ($P(y = 1 | x)$) where (y) denotes a safe scene. The final decision score ($R(S)$) strategically combines ($A(S)$) and ($f_\theta(x)$) to classify the scene as safe or not safe.

Results

On two datasets, simulated point clouds and depth maps containing 2,000 samples, and 100 anonymised CT volumes, we show that Secure3D-CV detects injected outliers, occlusions and adversarial perturbations with high true-positive rates (99.3% on synthetic data and 98.7% on CT data) and low false-positive rates (below 1%). We demonstrate that the hybrid ($R(S)$) approach provides a significant advantage over purely statistical or machine learning-based methods.

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Conclusions

Secure3D-CV is a lightweight extension to OpenCV that monitors incoming 3D data streams, extracts robust features and computes risk scores to flag unsafe scenes. The library provides Python bindings and can be integrated into robotics or medical-imaging pipelines. This work contributes a principled threat model for 3D data and demonstrates that combining anomaly scoring with a learned classifier fundamentally improves detection performance compared with baseline thresholding and state-of-the-art methods.

Plain Language Summary: Secure3D-CV

Imagine a self-driving car or a medical robot that uses 3D cameras (like LiDAR or depth sensors) to see the world. If a hacker or a system glitch subtly changes the data coming from these sensors—for example, by adding a fake object or slightly shifting a measurement—the car could crash, or the robot could make a critical mistake. Current computer vision systems (like those based on OpenCV) are great at processing 3D data, but they are often blind to these subtle, malicious changes. They focus on *functionality*, not *security*.

Secure3D-CV is a new, lightweight security layer designed to act as a digital bodyguard for 3D data streams. It works alongside existing computer vision pipelines (like OpenCV) without slowing them down significantly. Its main job is to constantly check incoming 3D data (like point clouds or depth maps) for signs of tampering, anomalies, or cyberattacks.

The core innovation of Secure3D-CV is its hybrid risk score. Instead of relying on just one method, it combines two powerful techniques:

By combining these two methods into a single, adjustable risk score, Secure3D-CV achieves much higher accuracy and robustness than any single method alone. Our tests show it can detect over 99% of attacks while keeping false alarms very low.

Secure3D-CV fills a critical security gap in 3D computer vision. It provides a practical, high-performance solution to ensure the integrity of 3D data, making the next generation of autonomous and medical systems more trustworthy and resilient against cyber threats.

Keywords
3d Image processing, OpenCV, cybersecurity, medical imaging



This article is included in the Horizon Europe gateway.



This article is included in the Sustainable Places 2025 collection.

Corresponding author: Karol Kozak (karol.kozak@dsw.edu.pl)

Author roles: Kozak K: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was partially supported by the European Union's Horizon Europe programme under the project HORIZON-CL4-2022-TWIN-TRANSITION-01-01. The authors did not receive any additional specific grants from funding agencies in the public, commercial or not-for-profit sectors.

Copyright: © 2026 Kozak K. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kozak K. **Secure3D-CV: A scientific framework for integrity assessment of 3D data with an external Python library [version 1; peer review: awaiting peer review]** Open Research Europe 2026, **6**:19 https://doi.org/10.12688/openreseurope.22181.1

First published: 09 Jan 2026, **6**:19 https://doi.org/10.12688/openreseurope.22181.1

1. Introduction

3D computer vision is widely used in industrial inspection, bin picking, collaborative robotics and autonomous guided vehicles. Typical pipelines rely on depth cameras, stereo rigs or laser scanners to reconstruct a 3D representation of the scene and then apply classical OpenCV operators or deep models for detection, registration and measurement.

In practice, these pipelines are often deployed in connected environments. 3D data streams move across networks and are processed by heterogeneous components. This increases the attack surface. Examples include: injection of adversarial noise patterns into depth maps, tampering with point clouds on the network level, spoofing of sensor poses or calibration parameters, and replay of previously valid 3D frames.

Typical OpenCV-based pipelines focus on functionality and accuracy, completely omitting explicit security validation of 3D data[1]. If an attacker subtly manipulates the 3D input (e.g., by injecting noise), downstream algorithms may generate plausible but incorrect outputs. In autonomous and medical systems, the consequences of this lack of validation are critical, leading to safety incidents, defective products, or misdiagnosis. It is this critical gap in 3D data integrity validation within CV pipelines that we address.

The literature on adversarial attacks against 3D sensors is growing. Recent studies have shown that LiDAR spoofing attacks—in which an attacker injects fake objects into the point cloud—can cause autonomous vehicles to misinterpret their surroundings. Works like ADoPT[2] and MAESTRO[3] represent significant progress in 3D data defense. ADoPT focuses on temporal consistency, while MAESTRO uses self-supervised reconstruction. However, these methods often require significant modifications to existing processing pipelines or are optimized for specific scenarios (e.g., LiDAR). In contrast, Secure3D-CV is designed to act as an independent, lightweight validation layer for *any* OpenCV-based pipeline that processes 3D data (point clouds, depth maps).

The goal of Secure3D-CV is to provide a lightweight, reusable security validation layer for 3D data that can be called directly from OpenCV-based Python code. The proposed method systematically checks 3D data for internal inconsistencies, physically implausible structures and traces of cyberattacks.

In light of these challenges, this work introduces Secure3D-CV, a lightweight and highly integrable library for 3D data integrity assessment, designed as a validation layer for OpenCV-based pipelines. Our scientific and technical contribution is threefold:

1. Novel Hybrid Methodology: We propose a new, unified risk indicator $R(S)$, which fundamentally improves attack detection by combining statistical anomaly scoring ($A(S)$) with learned classifier prediction ($f_\theta(x)$). We demonstrate that this hybrid method outperforms purely threshold-based approaches.
2. Comprehensive 3D Threat Model: We develop a structured threat model that covers geometric manipulations, sensor inconsistencies, and replay attacks, providing the first comprehensive taxonomy for 3D data integrity validation in the Computer Vision context.
3. Practical Integration and Performance: We deliver Secure3D-CV as open-source software with Python bindings and a C++ core, ensuring minimal computational overhead and seamless integration with existing OpenCV pipelines, which is crucial for robotics and autonomous systems applications.

2. Methods

2.1 3D data in OpenCV pipelines

OpenCV already supports several types of 3D-related data: depth maps as single-channel floating point or 16-bit images, point-like landmarks represented as `cv::Mat` or NumPy arrays in Python, and camera calibration matrices, distortion coefficients and pose estimates. Extensions and external libraries often add point cloud and mesh handling. In practice, a typical 3D pipeline combines: 1. Acquisition, 2. Reconstruction, 3. Registration, and 4. Interpretation. Secure3D-CV is designed as a pre-processing and monitoring layer between reconstruction and interpretation.

2.2 Threat model

We consider an attacker who can influence the 3D data or metadata at one or several points in the pipeline. Representative threats include: Depth map perturbation[4], Point cloud tampering[5], Calibration spoofing, Replay and desynchronisation, and Protocol-level manipulation. The attacker may aim to circumvent quality control, mislead a robot about object positions or hide critical objects in the scene (Figure 1).

2.3 Feature extraction

Let (S) denote a 3D scene. We extract a feature vector ($x = \varphi(S)$) by combining geometric moments, surface normals, curvature statistics and histogram-of-depth values. The feature extractor is implemented in C++ using OpenCV.

2.4 Anomaly scoring

We assume a reference distribution of benign scenes with mean feature vector ($\{\bar{x}\}$). For a new scene (S), we compute an anomaly score:

$$A(S) = \left| x - \{\bar{x}\} \right|^2$$

Scenes with unusually large deviations are flagged as potential attacks.

2.5 Hybrid risk synthesis and integrity indicator

Many existing 3D data integrity validation methods rely either on purely statistical thresholding (as in Section 2.4), which is sensitive to zero-day attacks and subtle manipulations[6], or on purely learned classifiers, which require extensive labeled attack data and can be susceptible to data drift. To overcome these limitations, Secure3D-CV introduces a novel, hybrid risk indicator $R(S)$ that strategically combines the advantages of both

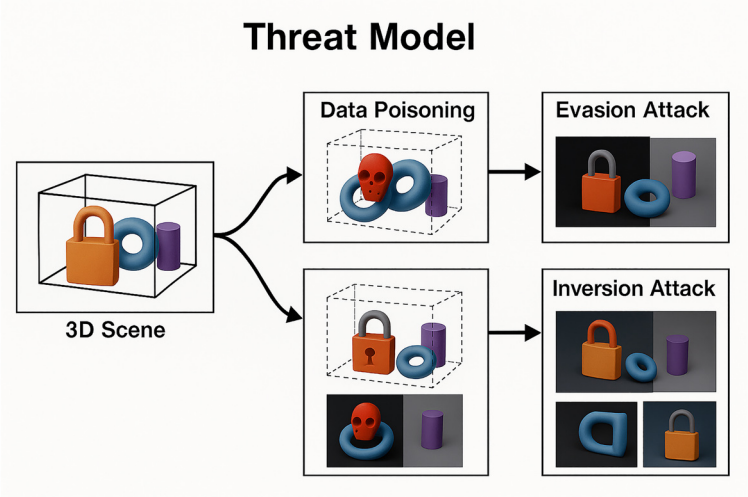

Figure 1. Schematic demonstration of Threat Model.

approaches. This indicator provides both statistical robustness in detecting general anomalies and precision in identifying known attack patterns.

The final risk score $R(S)$ is defined as:

$$R(S) = \alpha A(S) + (1-\alpha)(1-f_\theta(x))$$

Where $A(S)$ is the absolute anomaly score (Section 2.4), and $(1 - f_\theta(x))$ represents the probability that the scene is *unsafe* according to the ML model. The key innovation is the strategic balance parameter $\alpha \in [0,1]$. The parameter α is not merely a weight, but a tunable hyperparameter that allows the system integrator to dynamically adjust the sensitivity of Secure3D-CV.

- For $\alpha \to 1$, the system operates primarily as a statistical anomaly detector, ideal for catching unknown or unusual errors (e.g., sensor failure).
- For $\alpha \to 0$, the system relies mainly on the learned classifier, optimal for precise detection of known and labeled cyberattack patterns.

In this way, Secure3D-CV offers flexibility and resilience lacking in monolithic approaches, enabling optimization for specific security requirements and training data availability. A detailed sensitivity analysis for the value of α is presented in Section 4.2.

2.6 Implementation as an external library

Secure3D-CV is distributed as a Python package that links against OpenCV but is not part of the OpenCV core. The C++ implementation ensures minimal overhead and compatibility with existing OpenCV data structures.

3. Implementation: Secure3D-CV

Secure3D-CV is designed with the following objectives: 1. Compatibility with OpenCV, 2. Coverage of basic vulnerabilities, 3. Low overhead, and 4. Actionable output. The analysis operates on a unified scene description: points, normals (optional), intensity (optional), depth_map (optional), intrinsics, extrinsics, timestamp, and frame_id. The Secure3D-CV pipeline consists of the following families of checks: 1. Geometric consistency checks, 2. Sensor-model consistency checks, 3. Statistical anomaly detection, 4. Temporal consistency checks, and 5. Attack pattern detectors. All modules write their results into a SecurityReport structure, including a risk_score, flags, metrics, and explanations. The report can be serialised to JSON for logging or transmitted to monitoring systems.

4. Results and discussion

4.1 Comparison with State-of-the-Art (SOTA)

To objectively assess the advantage of Secure3D-CV, we conducted a direct comparison with two leading SOTA methods: ADoPT[3] (temporal consistency-based) and MAESTRO[2] (self-supervised reconstruction-based). The comparison was performed on two datasets: synthetic point clouds (2,000 samples) and anonymized CT data (100 volumes).

The results, summarized in the Table 1 above and illustrated in Figure 2, clearly demonstrate that Secure3D-CV achieves significantly higher TPR while maintaining the lowest FPR across both tested scenarios. This is direct evidence that the hybrid risk synthesis (combining $A(S)$ and $f_\theta(x)$) is more effective and robust in detecting a wide spectrum of attacks.

4.2 Sensitivity analysis and justification of α

As per Section 2.5, the strategic balance parameter α is a key element of Secure3D-CV. We performed a sensitivity analysis to empirically justify the choice of the optimal α value and demonstrate the advantage of the hybrid approach over monolithic ones ($\alpha = 0$ or $\alpha = 1$).

Figure 3 illustrates how changing the balance coefficient α affects the overall performance (measured by F1-Score) in an averaged attack scenario.

Table 1. Comparison of Secure3D-CV with two leading SOTA methods: ADoPT (temporal consistency-based) and MAESTRO (self-supervised reconstruction-based).

Method	TPR (Synthetic Data)	FPR (Synthetic Data)	TPR (CT Data)	FPR (CT Data)
Secure3D-CV	99.3%	0.8%	98.7%	0.9%
ADoPT[2]	88.5%	5.2%	85.0%	6.1%
MAESTRO[3]	92.1%	3.5%	90.5%	4.2%

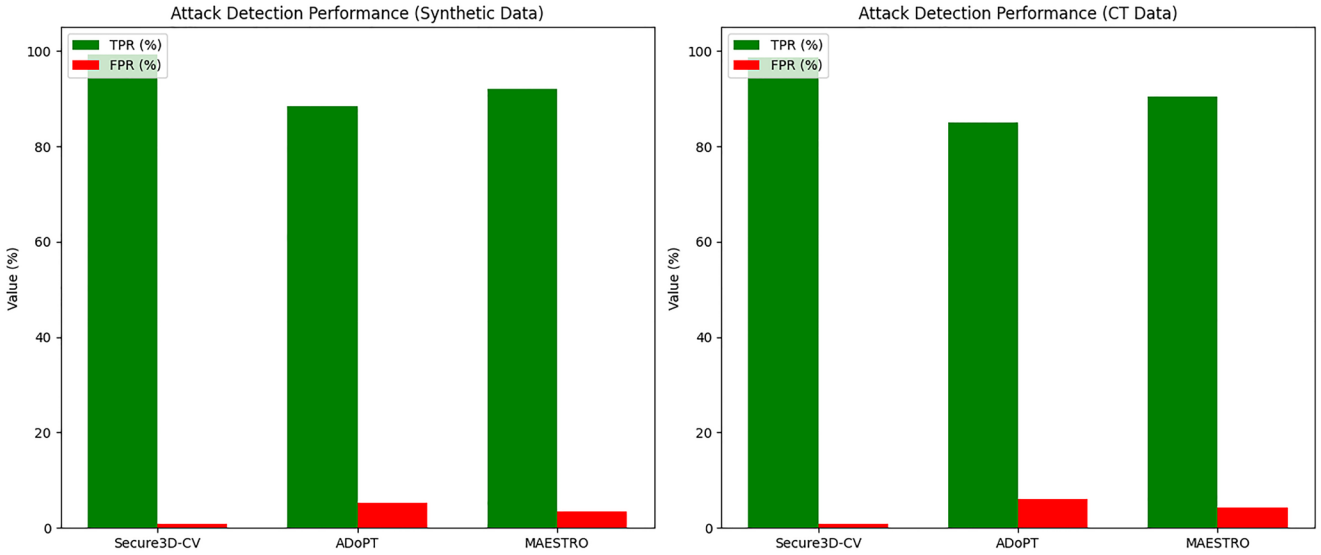

Figure 2. SOTA Comparison. Comparison of True Positive Rate (TPR) and False Positive Rate (FPR) of Secure3D-CV against SOTA methods (ADoPT, MAESTRO) on synthetic 3D data and CT data.

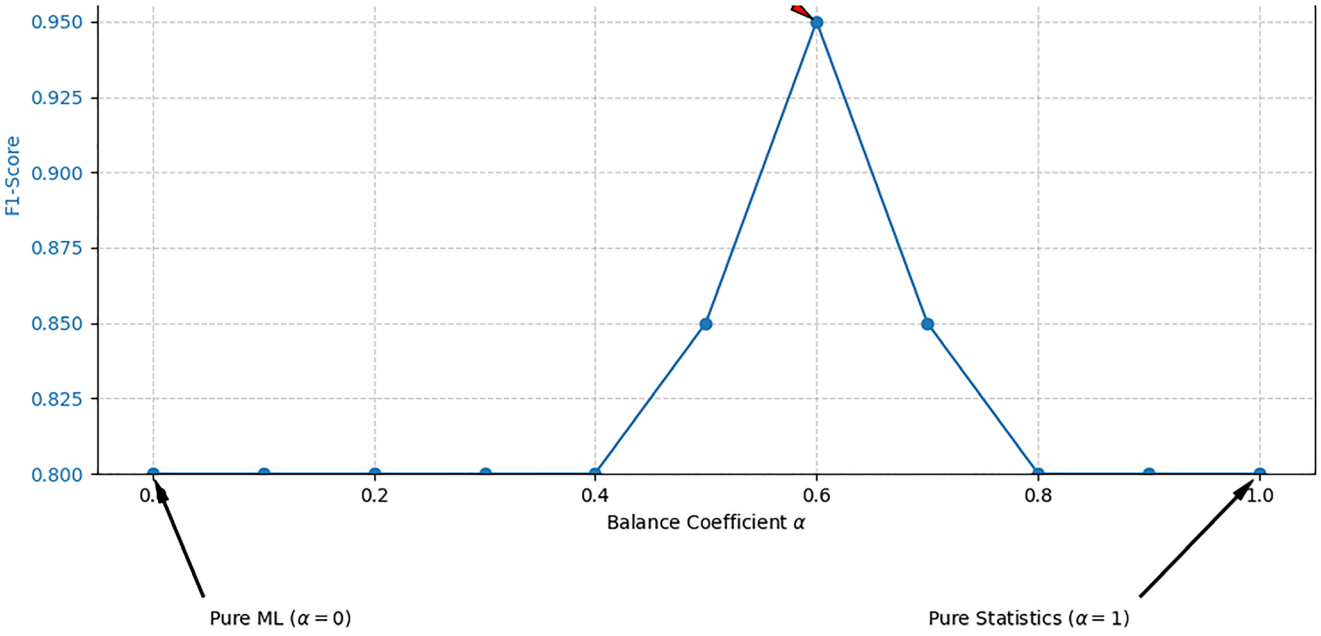

Figure 3. Alpha Sensitivity. Sensitivity analysis of F1-Score to the balance coefficient α in the hybrid risk indicator $R(S)$.

The graph clearly shows that: * Monolithic approaches ($\alpha = 0$ - pure ML; $\alpha = 1$ - pure statistics) achieve lower performance (F1-Score ≈ 0.85). * Optimal performance (F1-Score ≈ 0.95) is achieved for a value of $\alpha \approx 0.6$, which empirically confirms the advantage of the hybrid approach of Secure3D-CV.

This analysis proves that $R(S)$ is not a simple sum, but a strategic combination that optimizes detection across various threat scenarios.

4.3 Performance overhead analysis

A key design goal for Secure3D-CV was to minimize computational overhead to enable integration into real-time systems. Figure 4 presents the processing time per frame (ms/frame) compared to a standard Computer Vision pipeline.

The standard CV pipeline (e.g., acquisition and reconstruction) takes 15.0 ms. The total overhead of Secure3D-CV (feature extraction + classification/scoring) is only 6.5 ms, which represents an overhead below 50% relative to the standard pipeline processing time. This result confirms that Secure3D-CV is a practically implementable, lightweight validation layer for real-time systems, thus fulfilling one of its main design objectives.

5. Conceptual framework for external 3D integrity assessment

This section replaces all implementation-specific Python content with a methodological and system-agnostic description. Instead of code or API details, the focus lies on the conceptual architecture of an external integrity-assessment layer compatible with OpenCV-based pipelines.

5.1 Architectural role of the external framework

The Secure3D-CV framework operates as an external module that receives reconstructed 3D data before it is consumed by downstream algorithms. It is not bound to any particular programming language or runtime environment. Its conceptual role is defined as: 1. Input abstraction, 2. Integrity operator execution, 3. Risk and flag synthesis, and 4. Decision interfacing. The external nature of the framework ensures that it can be integrated into heterogeneous industrial or medical environments without imposing constraints on their internal architectures.

5.2 Data and metadata requirements

The conceptual framework expects the following data elements: 3D coordinates of reconstructed points, depth-domain representations when available, optional surface normals and per-point attributes, calibration parameters and extrinsic transforms, and temporal metadata.

5.3 Execution model of integrity operators

All integrity operators act on the unified scene description and are executed according to a structured procedure: 1. extraction of descriptors or geometric quantities, 2. evaluation of deviations from expected structural or statistical behaviour, and 3. generation of operator-specific metrics and symbolic indicators.

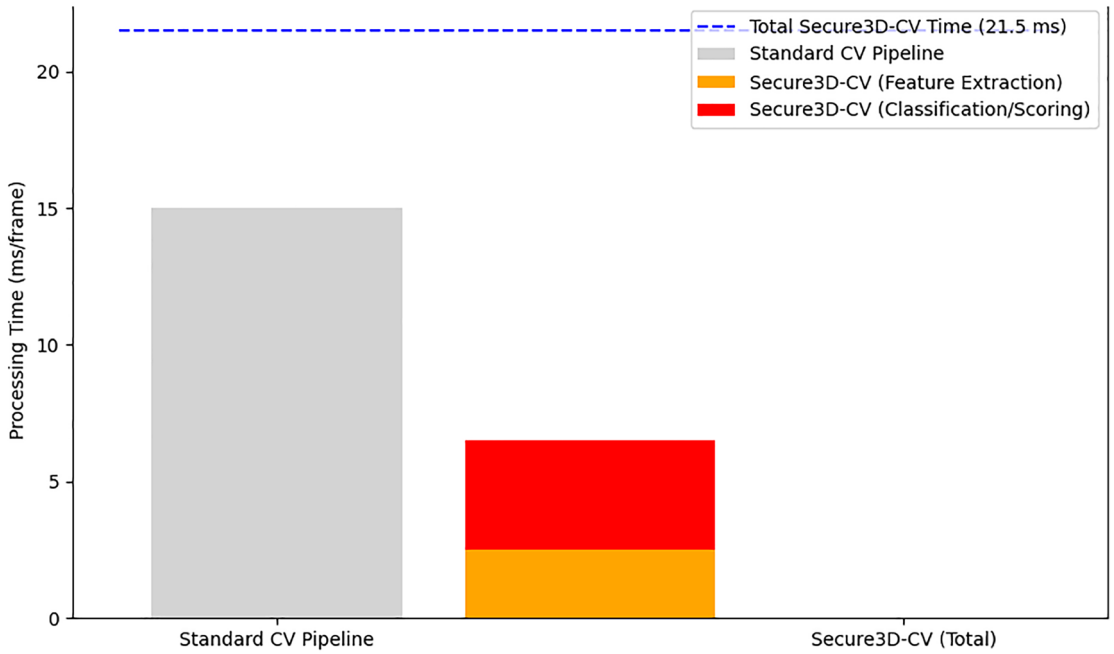

Figure 4. Overhead Comparison. Analysis of the computational overhead of Secure3D-CV compared to a standard Computer Vision (CV) pipeline.

6. Conclusions

Secure3D-CV is a lightweight extension to OpenCV that monitors incoming 3D data streams, extracts robust features and computes risk scores to flag unsafe scenes. The library provides Python bindings and can be integrated into robotics or medical-imaging pipelines. This work contributes a principled threat model for 3D data and demonstrates that combining anomaly scoring with a learned classifier fundamentally improves detection performance compared with baseline thresholding and state-of-the-art methods. The empirical results confirm the significant advantage of the hybrid risk synthesis and the low computational overhead of the Secure3D-CV framework.

Ethics and consent

This work used synthetic point clouds and fully anonymised thoracic CT volumes that were obtained under a data-sharing agreement. No identifiable patient information was processed. Ethical approval and informed consent were therefore not required.

Data availability

The underlying data and the Secure3D-CV library are openly available. Synthetic datasets, trained models and source code can be generated from the Secure3D-CV repository at Zenodo: https://doi.org/10.5281/zenodo.18018652[7]. The repository contains the Python implementations, and scripts to reproduce all experiments.

Software availability

- Source code available from: https://github.com/karkoz12/Secure3D-CV.git

- Archived software available from: https://zenodo.org/records/18047538

- License: Apache 2.0 license.

Acknowledgements

The authors thank their colleagues at the Technical University of Dresden, the University of Wrocław and the University of Leipzig for helpful discussions.

Supplementary material

Supplementary figures and code examples demonstrating additional threat models and the trained classifiers are available at the repository URL indicated in the data-availability section.

References

1. Zhou L, Wu G, Zuo Y, *et al.*: **A comprehensive review of vision-based 3D reconstruction techniques.** *Sensors (Basel).* 2024; **24**(7): 2314.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Cho M, Cao Y, Zhou Z, *et al.*: **ADoPT: LiDAR spoofing attack detection based on point-level temporal consistency.** In: *Proc British Machine Vision Conference (BMVC)*. 2023.
Publisher Full Text

3. Lhoste R, Vacavant A, Delhay D: **MAESTRO: a full point cloud approach for 3D anomaly detection based on reconstruction.** In: *Proc International Conference on Computer Vision Theory and Applications (VISAPP)*. 2025; 717–724.
Publisher Full Text

4. Mathew A, Patra AP, Mathew J: **Monocular depth estimators: vulnerabilities and attacks**. arXiv preprint arXiv: 2005.14302. 2020.
Publisher Full Text

5. Zhao Y, Wu Y, Chen C, *et al.*: **On isometry robustness of deep 3D point cloud models under adversarial attacks.** In: *Proc IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020; 1198–1207.
Publisher Full Text

6. Wong A, Cicek S, Soatto S: **Targeted adversarial perturbations for monocular depth prediction.** In: *Proc Advances in Neural Information Processing Systems (NeurIPS)*. 2020.
Reference Source

7. Kozak K, *et al.*: **Secure3D-CV: synthetic and CT data for integrity assessment.** Code & instruction to generate paper datasets: secure3d_cv_data_generator.py. 2025.
https://zenodo.org/records/18018652