UNINORM: NORMAL-DIRECTION SURFACE-DEVIATION MEASUREMENT FOR MULTI-CATEGORY 3D ANOMALY DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-category 3D anomaly detection must detect and localize defects across heterogeneous categories using only normal training data. Existing methods largely improve representations of normality while inheriting feature distance, reconstruction error, or field residual as anomaly evidence, coupling representation quality with the suitability of the resulting measurement. Under matched registered inputs and identical object pooling, we compare geometric residuals with learned implicit-field residuals and find performance differences. We therefore introduce UniNorm, which formulates template-aligned 3D anomaly detection as normal-direction surface-deviation measurement. UniNorm uses absolute normal projection to retain surface-normal departure while suppressing tangential correspondence motion, and optionally applies a nonparametric normal-dispersion correction (NULB) together with a shared contrast-tail object readout. The scorer has no learned parameters; only a lightweight category-specific nominal template is stored. On Anomaly-ShapeNet and Real3D-AD, UniNorm achieves 92.60/90.53% and 85.14/83.48% object-/point-level AUROC, respectively. With registration and object pooling matched, absolute normal projection improves over a class-conditioned implicit-field residual by 6.83/8.95 and 3.16/8.26 O/P-AUROC points on the two benchmarks. These results highlight anomaly measurement as a first-class design choice in template-aligned 3D anomaly detection.

1 INTRODUCTION

Three-dimensional anomaly detection (3D-AD) is an important component of automated quality inspection. Given only anomaly-free training samples, a detector must identify defective objects and localize abnormal surface regions at test time (Bergmann et al., 2022; Liu et al., 2023; Li et al., 2024). Point clouds expose geometric defects such as dents, bulges, holes, and missing parts more directly than images, but their unordered and sparse sampling also makes subtle defects difficult to distinguish from variations in density, coverage, correspondence, and pose (Ye et al., 2025; Zheng et al., 2025). Practical inspection systems must further handle many product categories, motivating multi-category detectors that avoid maintaining an independent learned model for every class (Cheng et al., 2025; Kim et al., 2026).

Most existing 3D-AD methods primarily improve *how normality is represented*. Feature-embedding approaches compare test descriptors with stored normal features (Horwitz & Hoshen, 2023; Wang et al., 2023; Zhao et al., 2026); reconstruction-based methods learn to restore normal points or features (Zhou et al., 2024c; Ye et al., 2025; Wu et al., 2026); and implicit methods represent nominal geometry using continuous fields (Zheng et al., 2025; Xiao et al., 2026). Recent multi-category methods further share geometry-aware encoders, reconstruction networks, or semantic priors across classes (Cheng et al., 2025; Kim et al., 2026; Liang et al., 2026). Despite these different representations, anomaly evidence is commonly inherited from the representation itself, through feature distance, reconstruction error, predicted offset, or field residual. Consequently, two distinct design choices are often coupled: *how normality is represented and how departure from normality is measured*.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

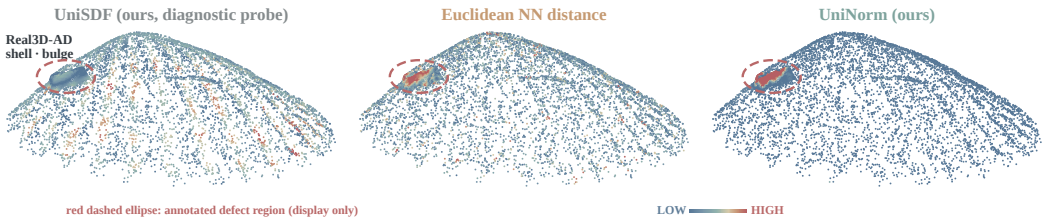

Figure 1: Anomaly-score comparison on an illustrative Real3D-AD shell scan. The implicit-field residual produces a broad response, and Euclidean distance responds to tangential correspondence motion. UniNorm suppresses the surface-wide baseline and retains a localized response near the defect. The dashed ellipse shows the annotation for visualization only; score maps are independently normalized.

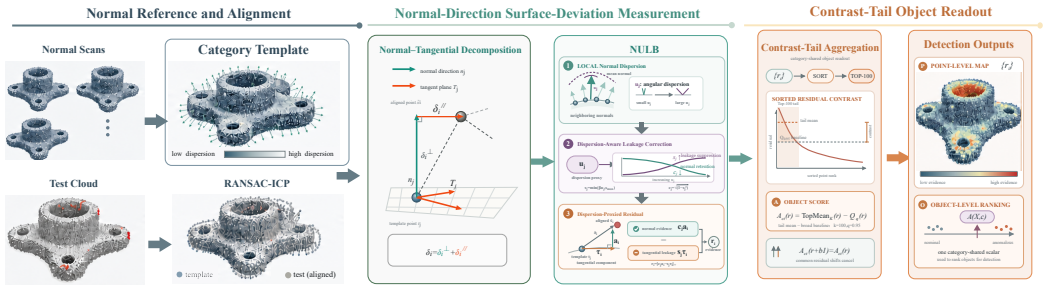

Figure 2: Overview of UniNorm. Normal training scans construct a category-specific lightweight template with a template-side normal-dispersion proxy, while RANSAC–ICP aligns each test cloud to that reference. The aligned displacement is decomposed into normal and tangential components. Absolute normal projection produces the point-level anomaly map, and the optional NULB dispersion-proxy correction refines it. The contrast-tail readout aggregates this map into the object score $A(X, c) = A_{\mathrm{ct}}(\mathbf{r}(X, c))$, with $K = 100$ and $q = 0.95$ in experiments.

This distinction is particularly important for geometric anomaly detection. A representation-native discrepancy does not necessarily provide a faithful measure of a physical surface defect. Point-to-point distance, for example, responds not only to displacement away from the nominal surface but also to tangential correspondence motion caused by sparse sampling or imperfect matching. Conversely, learned reconstruction or implicit-field residuals can produce spatially broad responses that do not closely follow localized surface defects. Figure 1 illustrates this behavior on the same registered object. These observations motivate treating *anomaly measurement itself as a first-class design choice*, rather than simply adopting the discrepancy naturally produced by a normality representation.

Based on this perspective, we propose UniNorm, which formulates template-aligned 3D anomaly detection as *normal-direction surface-deviation measurement*. A standard RANSAC–ICP front end aligns each test cloud with a nominal template constructed from normal training data (Fischler & Bolles, 1981; Besl & McKay, 1992). UniNorm then measures local displacement along the reference surface-normal direction, retaining geometric departure from the nominal surface while suppressing tangential correspondence motion. Because estimated surface normals may be unreliable in highly curved or sparsely sampled regions, we additionally introduce normal-uncertainty leakage balancing (NULB), a lightweight nonparametric correction based on template-side normal dispersion. A fixed contrast-tail readout converts the point-level evidence into an object-level anomaly score while reducing sensitivity to broad residual offsets. As illustrated in Fig. 2, category identity selects only a lightweight nominal template; the measurement rule, object-level readout, and hyperparameters are shared across categories, and the scorer contains no learned parameters.

A central goal of our experiments is not only to compare UniNorm with existing detectors, but also to examine the roles of anomaly measurement and representation design. We compare residuals with registered inputs and object-level aggregation held fixed. Geometric residuals additionally share

the same template and correspondences, while comparisons with SDF probes also involve different representations. Across Anomaly-ShapeNet and Real3D-AD, these experiments show that normal-direction surface deviation improves over a matched class-conditioned implicit-field residual, while the additional dispersion correction provides a smaller secondary refinement. Robustness, efficiency, and failure-mode analyses further characterize the operating regime of the proposed measurement.

Contributions.

- **A controlled finding on anomaly measurement.** Under matched registered inputs and identical object pooling, we compare geometric and learned implicit-field residuals. Within a fixed template representation, residual-component ablations isolate the effect of normal and tangential evidence.
- **A normal-direction formulation for template-aligned 3D-AD.** We formulate anomaly evidence as normal-direction surface-deviation measurement and instantiate it with absolute normal projection, which emphasizes surface-normal departure while suppressing tangential correspondence motion.
- **A lightweight shared scoring pipeline with explicit operating boundaries.** We complement the primary measurement with an optional nonparametric normal-dispersion correction and a category-shared contrast-tail readout. The scorer has zero learned parameters. We report residual and reference ablations, robustness, efficiency, and failure-case analyses.

2 Related Work

Normality representation for 3D anomaly detection. Existing 3D anomaly detection methods differ primarily in how they model normal geometry. Feature-based approaches represent normal objects using local or global descriptors and detect anomalies through deviations from stored normal features (Roth et al., 2022; Bergmann & Sattlegger, 2023; Horwitz & Hoshen, 2023; Wang et al., 2023; Zhao et al., 2026). Simple3D combines handcrafted geometric descriptors, spatial feature aggregation, and a normal prototype set (Cheng et al., 2026). Reconstruction-based methods instead learn to recover normal points or geometric structures, treating reconstruction discrepancy as anomaly evidence (Zhou et al., 2024c; Ye et al., 2025; Wu et al., 2026). Implicit approaches model nominal surfaces continuously and detect anomalies from deviations relative to the learned field (Park et al., 2019; Zheng et al., 2025; Xiao et al., 2026). While these paradigms improve the representation of normality, their anomaly scores are typically tied to the discrepancy naturally induced by the representation. Our work is complementary: rather than introducing a higher-capacity normality representation, we study how anomaly evidence should be measured once a nominal geometric reference is available.

Multi-category 3D anomaly detection. A growing line of work seeks to replace category-specific detectors with models that share representations across heterogeneous object classes. Unified anomaly detection was first explored through shared models across categories (You et al., 2022), while recent 3D approaches develop geometry-aware reconstruction, semantic disentanglement, and multi-scale feature fusion for multi-category detection (Cheng et al., 2025; Kim et al., 2026; Liang et al., 2026). Related zero-shot methods instead pursue category or semantic generalization using vision-language priors (Zhou et al., 2024a;b). UniNorm addresses a different, closed-set setting: the category label is known at inference and selects only a lightweight nominal template, whereas the anomaly measurement, object-level readout, and hyperparameters are shared across all categories. Our goal is therefore not to learn a single geometric representation for unrelated object classes, but to determine whether a shared measurement rule can operate effectively over category-specific geometric references.

Surface inspection, geometric deviation, and direction-aware evidence. Measuring geometric deviation from a nominal surface has a long history in dimensional inspection and free-form surface metrology. Classical CAD-directed inspection aligns measured samples with a nominal design and evaluates their geometric discrepancy from the reference surface (Menq et al., 1992), while subsequent surveys formalize surface localization, measurement-to-design comparison, and surface-to-surface distance evaluation as central components of free-form inspection (Li & Gu, 2004; Savio et al., 2007). Closely related geometric principles also appear in 3D registration: point-to-plane

registration minimizes displacement along the target surface normal rather than full point-to-point distance (Chen & Medioni, 1992), and later work studies point-to-surface alignment through the geometry of the squared distance function (Pottmann et al., 2004). These methods establish normal-direction deviation as a useful geometric quantity, but they primarily address registration, dimensional verification, or tolerance inspection rather than unsupervised anomaly detection.

Recent 3D-AD methods increasingly exploit related geometric cues. PO3AD predicts point offsets toward normal geometry (Ye et al., 2025), geometry-aligned and anomaly-aware reconstruction incorporates geometric consistency into the reconstruction process (Wu et al., 2026), and registration itself has been shown to provide robustness to pose variation (Yu et al., 2025). In these approaches, however, the final anomaly evidence remains coupled to learned offsets, reconstruction errors, or other representation-dependent residuals. Building on the distinction between representation and measurement, Sec. 3 introduces UniNorm, which repurposes normal-direction surface deviation as the anomaly measurement itself while explicitly accounting for local normal unreliability and category-shared object scoring.

3 METHOD

UniNorm separates the *normality reference* from the *anomaly measurement*. For each known category, normal training data provide a lightweight nominal surface reference, while the same measurement rule and object-level readout are used across all categories. As illustrated in Fig. 2, inference consists of three stages: (1) align the test point cloud with the category-specific nominal template; (2) measure local surface deviation primarily along the template normal, with an optional correction for unreliable normals; and (3) aggregate the resulting point-level evidence into a category-shared object score. The scoring pipeline contains no learned parameters.

3.1 PROBLEM SETTING AND OVERVIEW

Let $\mathcal{C}$ denote the set of known product categories. For each category $c \in \mathcal{C}$, the training set $\mathcal{X}_c$ contains only normal point clouds. We consider the standard closed-set setting in which the category label is available at test time and is used only to select the corresponding nominal reference. Given a test point cloud $X = \{x_i\}_{i=1}^{M}$ from category c, the detector returns point-level anomaly evidence $\mathbf{r}(X, c) = (r_i)_{i=1}^{M}$ and an object-level anomaly score $A(X, c)$.

Each category is represented by a lightweight template

$$T_c = \{(t_j, n_j, u_j)\}_{j=1}^{N_c}, \qquad \|n_j\|_2 = 1, \tag{1}$$

where t_j is a nominal surface point, n_j is its unit surface normal, and u_j is a template-side normal-dispersion proxy introduced in Sec. 3.4. The template may be sampled from a normal mesh or constructed from normal training scans. Importantly, the template is the only category-specific state: the registration procedure, point-level measurement, object-level readout, and scoring hyperparameters are shared across categories.

3.2 NOMINAL REFERENCE AND REGISTRATION

Template normals are computed from adjacent mesh faces or local surface neighborhoods. For deterministic preprocessing, we orient each normal away from the template centroid $\bar{t}$:

$$n_j \leftarrow \begin{cases} n_j, & \langle n_j, t_j - \bar{t} \rangle \geq 0, \\ -n_j, & \langle n_j, t_j - \bar{t} \rangle < 0. \end{cases} \tag{2}$$

The anomaly measurements introduced below are invariant to the sign of the normal; this orientation therefore serves only to make preprocessing and visualization consistent.

For a test cloud X, we estimate a rigid transform using coarse RANSAC followed by ICP (Fischler & Bolles, 1981; Besl & McKay, 1992):

$$\begin{aligned} \widetilde{X} &= G_{X,c}(X), \\ G_{X,c} &\approx \arg\min_{G \in SE(3)} d_{\mathrm{CD}}(G(X), T_c)\,, \end{aligned} \tag{3}$$

where $G_{X,c}$ is the sample-specific rigid transformation and d_{CD} denotes the alignment objective. All subsequent anomaly measurements operate on the registered cloud $\widetilde{X} = \{\widetilde{x}_i\}_{i=1}^{M}$.

3.3 Normal-Direction Surface-Deviation Measurement

For each aligned test point $\widetilde{x}_i$, we identify its nearest template point and define the correspondence displacement as

$$j(i) = \arg\min_j \|\widetilde{x}_i - t_j\|_2, \qquad \delta_i = \widetilde{x}_i - t_{j(i)}. \tag{4}$$

The displacement can be decomposed in the local surface coordinate system into normal and tangential components:

$$\delta_i = \underbrace{\langle \delta_i, n_{j(i)} \rangle n_{j(i)}}_{\delta_i^{\perp}} + \underbrace{\delta_i - \langle \delta_i, n_{j(i)} \rangle n_{j(i)}}_{\delta_i^{\parallel}}. \tag{5}$$

Here, $\delta_i^{\perp}$ measures departure from the nominal surface along its local normal, whereas $\delta_i^{\parallel}$ represents in-plane correspondence motion. Euclidean nearest-neighbor distance combines both components. This can be undesirable for sparsely sampled point clouds, where different samplings of the same surface may exhibit large tangential displacement even in the absence of a defect.

We therefore define the primary anomaly evidence using the absolute normal projection,

$$a_i = \left| \langle \delta_i, n_{j(i)} \rangle \right|, \qquad \tau_i = \left\| \delta_i^{\parallel} \right\|_2. \tag{6}$$

The quantity a_i measures normal-direction surface deviation, while τ_i quantifies tangential mismatch. The absolute value treats inward and outward surface deviations symmetrically. Using a_i directly defines the core UniNorm measurement. In the experiments, we refer to this uncorrected variant as UniNorm-Abs only when it is necessary to distinguish it from the optional correction introduced next.

3.4 Normal-Uncertainty Leakage Balancing

Absolute normal projection assumes that the matched template normal reliably represents the local nominal surface. In practice, estimated normals may vary in curved, sparsely sampled, or geometrically ambiguous regions. In such regions, tangential correspondence motion can leak into the measured normal component and produce a spurious anomaly response.

We therefore introduce *normal-uncertainty leakage balancing* (NULB) as an optional, nonparametric correction. Here, "uncertainty" denotes a template-side dispersion proxy rather than a calibrated probabilistic uncertainty. The correction estimates how much of the observed normal response could plausibly arise from local normal variation and discounts the corresponding tangential leakage.

For brevity write $j = j(i)$. Let $n_j^{\star}$ denote the unknown ideal normal and let α_j be its acute angular deviation from the estimated normal n_j. Writing $n_j^{\star} = \cos\alpha_j\, n_j + \sin\alpha_j\, v_j$ for a unit tangent direction v_j, the reverse triangle inequality gives

$$\left| \langle \delta_i, n_j^{\star} \rangle \right| \geq \cos\alpha_j\, a_i - \sin\alpha_j\, \tau_i. \tag{7}$$

This relation highlights an asymmetry: for small angular error, attenuation of the normal signal enters through the cosine term, whereas tangential leakage enters through the sine term. We estimate the potential leakage using the sign-invariant RMS sine dispersion of the K_n neighboring template normals:

$$u_j = \sqrt{\frac{1}{K_n} \sum_{k \in \mathcal{N}_{K_n}(j)} \left(1 - |n_j^{\mathsf{T}} n_k|^2\right)}. \tag{8}$$

The quantity u_j is computed once from normal training geometry and does not depend on test samples or anomaly labels. We map this dispersion proxy to a leakage coefficient s_j and a normal-retention coefficient c_j:

$$s_j = \min(\beta u_j, s_{\max}), \qquad c_j = \sqrt{1 - s_j^2}. \tag{9}$$

The corrected point residual is then

$$r_i = \left[c_{j(i)} a_i - s_{j(i)} \tau_i\right]_+, \qquad [z]_+ = \max(z, 0). \tag{10}$$

If $s_j \geq \sin \alpha_j$, Eq. (7) yields a conditional lower bound. In practice, s_j is a fixed surrogate rather than a calibrated upper confidence bound on $\sin \alpha_j$. NULB should therefore be interpreted as a structured dispersion-based correction, rather than a probabilistic uncertainty estimate. Setting $\beta = 0$ recovers the uncorrected absolute normal projection.

3.5 Contrast-Tail Object Readout

Point-level defects are typically spatially sparse. A global mean can therefore dilute their response, whereas a maximum is highly sensitive to isolated outliers. Let $\mathbf{r} = (r_i)_{i=1}^M$ and let $r_{(1)} \geq \cdots \geq r_{(M)}$ denote its entries sorted in descending order and define

$$\mathrm{TopMean}_K(\mathbf{r}) = \frac{1}{K} \sum_{i=1}^{K} r_{(i)}. \tag{11}$$

UniNorm defines the contrast-tail readout

$$A_{\mathrm{ct}}(\mathbf{r}) = \mathrm{TopMean}_K(\mathbf{r}) - Q_q(\mathbf{r}), \tag{12}$$

where $Q_q(\mathbf{r})$ is the within-cloud q-quantile. The object score is $A(X, c) = A_{\mathrm{ct}}(\mathbf{r}(X, c))$. The first term summarizes sparse high-residual evidence, while the second estimates the broad residual baseline of the same test object. This subtraction makes the score invariant to a sample-wide additive offset:

$$A_{\mathrm{ct}}(\mathbf{r} + b\mathbf{1}) = A_{\mathrm{ct}}(\mathbf{r}). \tag{13}$$

This property allows the same object-level readout to be applied across categories with different residual baselines. It does not, however, remove spatially varying misregistration or missing surface coverage; these cases are examined explicitly in Sec. 4 and the supplementary analyses. Algorithm 1 summarizes the full inference procedure.

4 Experiments

4.1 Experimental Setup

Datasets and metrics. We use Anomaly-ShapeNet (40 synthetic categories) (Li et al., 2024) and Real3D-AD (12 real-scan categories) (Liu et al., 2023). Metrics are category-macro O-AUROC/O-AP and point-level P-AUROC (points pooled within each category, then macro-averaged). Paired bootstrap intervals are over categories.

Reference construction and registration. Templates come only from official normal training objects (no test object is used); one template per category (52 total) is capped at 20k points and shared by all residual variants. RANSAC–ICP is identical for all variants in our matched comparisons. Provenance and alternative references are in Appendix C.

Scoring configuration. Core residual is absolute normal projection. Optional NULB: K_n=30, β=0.5, $s_{\max}$=0.95. The global β was chosen exploratorily using labeled Anomaly-ShapeNet test metrics and then frozen for Real3D-AD (Appendix D). Contrast-tail readout: K=100, q=0.95 (fixed in advance). Results use fixed protocol-v2 registered caches.

Matched protocol. Published baselines are for overall benchmarking only. For comparisons with registration and object pooling controlled, we train a class-conditioned UniSDF probe and a category-specific PASDF-style SDF bank using normal training objects (UniSDF: 8×768, 64-D class embedding; bank width 256), then query both at the same registered test points. All residuals in Table 2 share contrast-tail pooling (K=100, q=0.95). Geometric residuals share one template; the SDF probes use learned representations. NULB is optional.

Table 1: Mean object-/point-level AUROC (O/P, %) on Anomaly-ShapeNet (40 classes) and Real3D-AD (12 classes). C denotes category-specific detectors, including learned models and nonparametric reference memories; U denotes unified/multi-category detectors with a shared model or scoring rule. UniNorm uses a shared scorer with a lightweight category-specific geometric reference. Published results are not component matched. Bold marks the best U entry in each column. ISMP does not report per-category object-level AUROC on Anomaly-ShapeNet.

Type	Method	Venue	AS O/P	R3D O/P
C	PASDF	ICCV'25	90.7/89.6	80.2/74.5
C	ISMP	AAAI'25	–/69.1	75.7/83.6
C	PO3AD	CVPR'25	83.9/89.9	76.6/65.0
C	CASL	AAAI'26	88.7/89.9	82.3/88.2
C	Simple3D	AAAI'26	86.0/92.9	80.4/92.3
U	MC3D-AD	IJCAI'25	83.6/75.9	78.2/76.8
U	MFF-M3AD	NN'26	89.9/77.1	81.1/78.9
U	SeDiR	CVPR'26	**93.3**/81.1	81.0/80.6
U	UniNorm (ours)	—	92.6/**90.5**	**85.1/83.5**

4.2 Comparison with State-of-the-Art Methods

Table 1 compares UniNorm with recent category-specific methods (PASDF (Zheng et al., 2025), ISMP (Liang et al., 2025), PO3AD (Ye et al., 2025), CASL (Zha et al., 2026), and Simple3D (Cheng et al., 2026)) and unified/multi-category methods (MC3D-AD (Cheng et al., 2025), MFF-M3AD (Liang et al., 2026), and SeDiR (Kim et al., 2026)). Because published baselines use their original preprocessing, representations, and scoring protocols, these results should be interpreted as overall benchmark comparisons rather than component-matched evidence. Among multi-category methods, UniNorm provides the strongest point-level localization on both benchmarks and the highest object-level AUROC on Real3D-AD, while SeDiR obtains the higher object-level AUROC on Anomaly-ShapeNet. Category-specific approaches can achieve stronger results in some settings and maintain category-specific learned models or reference memories; Simple3D uses nonparametric prototype memory. Full per-category comparisons are reported in Appendix A.

4.3 Controlled Analysis of Anomaly Measurement

4.3.1 Matched Measurement Comparison

To compare geometric and implicit-field residuals under matched registration and object pooling, we use two diagnostic probes: a bank of category-specific PASDF-style SDFs and a class-conditioned UniSDF with a shared implicit field. Both are queried at exactly the same registered test points as the geometric residuals. In Table 2, every row further uses the same contrast-tail object readout ($K = 100$, $q = 0.95$), giving a comparison with registration and aggregation fixed; the underlying representations are not held fixed. Neither SDF probe belongs to the final UniNorm inference pipeline.

Under this protocol, replacing the class-conditioned SDF residual with absolute normal projection improves O/P-AUROC by **6.83/8.95** points on Anomaly-ShapeNet and **3.16/8.26** points on Real3D-AD. The category-specific SDF bank improves over the shared UniSDF on several metrics, but remains below normal-direction measurement under the same readout, particularly for point localization. This indicates that the observed gap cannot be explained solely by whether the implicit representation is category-specific or shared.

Against matched UniSDF, absolute normal projection has higher object- and point-level AUROC in 27 and 39 of the 40 Anomaly-ShapeNet categories and in 10 and 10 of the 12 Real3D-AD categories; seven Anomaly-ShapeNet object-level comparisons are ties. With 10,000 paired category-bootstrap draws (seed 2026), the displayed object-level mean gaps are 6.83 and 3.16 points. Bootstrap intervals computed from unrounded per-category scores are $[3.97, 10.08]$ and $[1.38, 5.19]$, respectively (percentile 95%).

Table 2: Residual comparison with matched registration and object pooling (%). All residuals use identical registered inputs and the same contrast-tail object readout ($K = 100$, $q = 0.95$). O, AP, and P denote O-AUROC, O-AP, and P-AUROC, respectively. P-AUROC depends only on the residual channel.

	Anomaly-ShapeNet			Real3D-AD		
Residual	O	AP	P	O	AP	P
Category-specific SDF	91.43	93.32	89.17	81.94	79.16	75.44
Class-conditioned SDF (UniSDF)	85.10	89.06	81.35	81.54	79.24	75.22
Euclidean NN	89.96	92.25	83.61	75.87	76.63	78.89
Absolute normal projection	91.93	93.53	90.30	84.70	81.76	**83.48**
UniNorm (+NULB)	**92.60**	**93.98**	**90.53**	**85.14**	**82.21**	**83.48**

Table 3: Residual-channel ablation with identical object pooling (%).

	Anomaly-ShapeNet		Real3D-AD	
Residual	O	P	O	P
Euclidean NN	89.96	83.61	75.87	78.89
Tangential	63.55	52.25	55.60	52.53
Normal + tangent	91.25	85.99	81.76	80.36
Absolute normal projection	91.93	90.30	84.70	**83.48**
UniNorm (+NULB)	**92.60**	**90.53**	**85.14**	**83.48**

4.3.2 Residual Component Analysis

Table 3 isolates the geometric content of the residual by changing only the point-level measurement while keeping the object readout fixed. Tangential displacement is near chance at the point level. Normal+tangent uses $a_i + \tau_i$, indicating that motion within the local tangent plane provides little reliable anomaly evidence under the evaluated template-aligned setting. Reintroducing the tangential component also degrades the normal-direction residual, while full Euclidean distance is weaker than absolute normal projection. These results support the geometric premise in Eq. (5): surface-normal departure carries the useful defect signal, whereas tangential correspondence motion is predominantly nuisance variation.

4.3.3 Dispersion-Proxy Refinement

NULB is designed as a secondary correction for regions where the estimated template normal is locally unstable. Relative to uncorrected absolute normal projection, it yields a paired O-AUROC gain of 0.67 points on Anomaly-ShapeNet (95% CI $[0.33, 1.04]$) and 0.44 points on Real3D-AD ($[0.09, 0.85]$). The corresponding P-AUROC changes are $+0.23$ ($[0.06, 0.42]$) and 0.00 ($[-0.25, 0.29]$), respectively. The same $\beta = 0.5$, explored on Anomaly-ShapeNet, is transferred unchanged to Real3D-AD. We report the fixed NULB variant as the headline configuration because it improves mean O-AUROC on both benchmarks, while its mean Real3D-AD P-AUROC is unchanged. It is not uniformly better for localization: for broken/thin-boundary defects, $\beta = 0$ (UniNorm-Abs) avoids the observed point-level loss. Parameter sensitivity and the controlled tangential-leakage simulation are reported in Appendix D.

4.3.4 Object-Level Readout Robustness

The contrast-tail readout is stable over a broad range of K with a single $K = 100$ used for all categories and both datasets. Learned fusion and pseudo-selected pooling do not beat this fixed statistic (Appendix D); residual $\times$ aggregation results in Appendix D.1 show that absolute normal projection remains competitive under raw Top-K and contrast-tail pooling.

4.4 Efficiency

UniNorm has zero learned scoring parameters. Across 52 categories, templates store 1.04M points (27.70 MiB). Post-registration scoring is 18.71–20.89 ms/scan on CPU; registration (509.5 ms/scan) remains the deployment bottleneck (Appendix F). Figure 3 summarizes the accuracy–efficiency trade-off.

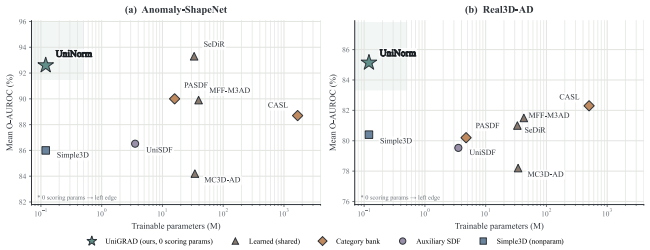

Figure 3: Mean O-AUROC versus learned scoring parameters on (a) Anomaly-ShapeNet and (b) Real3D-AD. UniNorm (star) stores category templates but has no learned scorer (zero shown at 10^{-1} on the log axis). Bank totals: PASDF 15.87M/4.76M and CASL 41.95M$\times$40/$\times$12 for ShapeNet/Real3D-AD. UniSDF uses auxiliary raw Top-500 pooling; pipelines are not component matched.

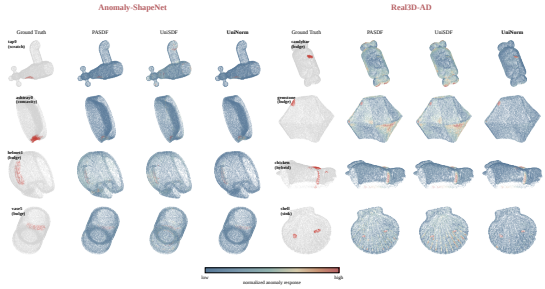

Figure 4: Eight-example point-level comparison (GT / PASDF-style SDF probe / UniSDF / UniNorm). Independent 5th–99.5th percentile color scale (display only). An enlarged view is in Appendix E.

4.5 Qualitative Behavior and Failure Modes

Figure 1 shows a Real3D-AD shell scan: the implicit-field residual is broad, Euclidean distance retains correspondence artifacts, and UniNorm localizes the defect. Figure 4 compares eight defects across both benchmarks; full-scale panels appear in Appendix E.

The failure cases delimit the method: NULB can reduce P-AUROC on broken/thin-boundary defects (-9.34 vs. UniNorm-Abs); one-way maps cannot score fully missing surface; failed registration should trigger reacquisition.

5 Discussion

With registration and pooling matched, normal-direction residuals outperform the evaluated SDF probes; fixed-template ablations support the benefit of normal over tangential evidence. Comparisons with learned fields also reflect representation differences. NULB is secondary. Normal-direction distance is an explicit anomaly-measurement operator, not a new primitive. Alignment remains the bottleneck. Settings are closed-set; NULB is a dispersion proxy (fallback: UniNorm-Abs, $\beta = 0$); one-way maps cannot score missing surface.

6 Conclusion

UniNorm measures anomaly evidence as normal-direction surface deviation with category-shared contrast-tail readout and no learned scoring parameters. Normal-direction residuals outperform the evaluated implicit-field probes under matched registration and pooling, while fixed-template ablations support the value of normal-direction evidence.

REPRODUCIBILITY STATEMENT

All primary results use fixed registered point-cloud caches, a common metric implementation, and only official normal-training objects for template and diagnostic-probe construction. The global NULB coefficient was selected after inspecting labeled Anomaly-ShapeNet test outcomes, making that benchmark a development set for β; Real3D-AD used the frozen coefficient. The main scoring hyperparameters (K_n, β, $s_{\max}$, K, and q) are specified in Sec. 4.1, shared across categories, and transferred unchanged from Anomaly-ShapeNet to Real3D-AD where stated. Matched residual comparisons use identical registered queries and identical object-level aggregation. An anonymized core-code release is available at `https://anonymous.4open.science/r/uninorm-code-AAD5/`, covering template registration, normal-direction residual measurement, NULB, and contrast-tail readout. The supplementary results archive provides anonymized configuration manifests and CSVs for the matched comparisons, NULB analyses, reference study, and robustness tests.

AI USE STATEMENT

During the preparation of this work, the authors used generative AI tools to assist with manuscript drafting and language editing, literature organization, methodological and experimental feedback, interpretation and presentation of results, LaTeX formatting, and figure-editing support. All AI-assisted outputs were independently reviewed and verified by the authors, who take full responsibility for the final manuscript.

REFERENCES

Paul Bergmann and David Sattlegger. Anomaly detection in 3D point clouds using deep geometric descriptors. In *WACV*, pp. 2613–2623, 2023.

Paul Bergmann, Xin Jin, David Sattlegger, and Carsten Steger. The MVTec 3D-AD dataset for unsupervised 3D anomaly detection and localization. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 202–213, 2022.

Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

Yang Chen and Gérard Medioni. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

Jiayi Cheng, Can Gao, Jie Zhou, Jiajun Wen, Tao Dai, and Jinbao Wang. MC3D-AD: A unified geometry-aware reconstruction model for multi-category 3D anomaly detection. In *IJCAI*, pp. 837–845, 2025. doi: 10.24963/ijcai.2025/94.

Yuqi Cheng, Yihan Sun, Hui Zhang, Weiming Shen, and Yunkang Cao. Towards high-resolution 3D anomaly detection: A scalable dataset and real-time framework for subtle industrial defects. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(5):3327–3334, 2026. doi: 10.1609/aaai.v40i5.37328.

Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24 (6):381–395, 1981.

Eliahu Horwitz and Yedid Hoshen. Back to the feature: Classical 3D features are (almost) all you need for 3D anomaly detection. In *CVPR*, pp. 2967–2976, 2023.

SuYeon Kim, Wongyu Lee, and MyeongAh Cho. A semantically disentangled unified model for multi-category 3D anomaly detection. In *CVPR*, pp. 33036–33045, 2026.

Wenqiao Li, Xiao Xu, Yao Gu, Bozhong Zheng, Shenghua Gao, and Ying Wu. Towards scalable 3D anomaly detection and localization: A benchmark via 3D anomaly synthesis and a self-supervised learning network. In *CVPR*, pp. 22207–22216, 2024.

Yadong Li and Peihua Gu. Free-form surface inspection techniques state of the art review. *Computer-Aided Design*, 36(13):1395–1417, 2004. doi: 10.1016/j.cad.2004.02.009.

Hanzhe Liang, Guoyang Xie, Chengbin Hou, Bingshu Wang, Can Gao, and Jinbao Wang. Look inside for more: Internal spatial modality perception for 3D anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(5):5146–5154, 2025. doi: 10.1609/aaai.v39i5.32546.

Hanzhe Liang, Chenxi Hu, Yejin Tang, Linlin Shen, Jinbao Wang, and Can Gao. MFF-M3AD: A unified reconstruction method with multi-scale feature fusion for multi-category 3D anomaly detection. *Neural Networks*, 203:109131, 2026. doi: 10.1016/j.neunet.2026.109131.

Jiaqi Liu, Guoyang Xie, Ruitao Chen, Xinpeng Li, Jinbao Wang, Yong Liu, Chengjie Wang, and Feng Zheng. Real3D-AD: A dataset of point cloud anomaly detection. *NeurIPS*, 36:30402–30415, 2023.

Chia-Hsiang Menq, Hong-Tzong Yau, and Gwan-Ywan Lai. Automated precision measurement of surface profile in cad-directed inspection. *IEEE Transactions on Robotics and Automation*, 8(2):268–278, 1992. doi: 10.1109/70.134279.

Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *CVPR*, pp. 165–174, 2019.

Helmut Pottmann, Stefan Leopoldseder, and Michael Hofer. Registration without icp. *Computer Vision and Image Understanding*, 95(1):54–71, 2004. doi: 10.1016/j.cviu.2004.04.002.

Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *CVPR*, pp. 14318–14328, 2022.

Enrico Savio, Leonardo De Chiffre, and Robert Schmitt. Metrology of freeform shaped parts. *CIRP Annals*, 56(2):810–835, 2007. doi: 10.1016/j.cirp.2007.10.008.

Yixuan Wang, Jingjing Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, and Chengjie Wang. Multimodal industrial anomaly detection via hybrid fusion. In *CVPR*, pp. 8032–8041, 2023.

Linchun Wu, Qin Zou, Yuanhao Yue, and Zhongyuan Wang. Geometry-aligned and anomaly-aware reconstruction for 3D anomaly detection. In *CVPR*, pp. 14648–14657, 2026.

Haibo Xiao, Hanzhe Liang, Jie Zhou, Jinbao Wang, and Can Gao. Learning discriminative signed distance functions from multi-scale level-of-detail features for 3D anomaly detection. *arXiv preprint arXiv:2605.03437*, 2026.

Jianan Ye, Weiguang Zhao, Xi Yang, Guangliang Cheng, and Kaizhu Huang. PO3AD: Predicting point offsets toward better 3D point cloud anomaly detection. In *CVPR*, pp. 1353–1362, 2025.

Zhiyuan You, Lei Cui, Yujun Shen, Kai Yang, Xin Lu, Yu Zheng, and Xinyi Le. A unified model for multi-class anomaly detection. In *NeurIPS*, volume 35, pp. 4571–4584, 2022.

Yuyang Yu, Zhengwei Chen, Xuemiao Xu, Lei Zhang, Haoxin Yang, Yongwei Nie, and Shengfeng He. Registration is a powerful rotation-invariance learner for 3D anomaly detection. In *NeurIPS*, volume 38, 2025.

Yaohua Zha, Yuerong Xue, Chunlin Fan, Yuansong Wang, Tao Dai, Ke Chen, and Shu-Tao Xia. CASL: Curvature-augmented self-supervised learning for 3D anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(15):12340–12348, 2026. doi: 10.1609/aaai.v40i15.38226.

Baozhu Zhao, Xiaohan Zhang, Jingfeng Guo, and Qi Liu. PointCore: An efficient framework for unsupervised point cloud anomaly detection using joint local-global features. *Neural Networks*, 197:108446, 2026. doi: 10.1016/j.neunet.2025.108446.

Bozhong Zheng, Jinye Gan, Xiaohao Xu, Xintao Chen, Wenqiao Li, Xiaonan Huang, Na Ni, and Yingna Wu. Bridging 3D anomaly localization and repair via high-quality continuous geometric representation. In *ICCV*, pp. 27063–27072, 2025.

Qihang Zhou, Guansong Pang, Yu Tian, Shibo He, and Jiming Chen. AnomalyCLIP: Object-agnostic prompt learning for zero-shot anomaly detection. In *ICLR*, 2024a.

Qihang Zhou, Jiangtao Yan, Shibo He, Wenchao Meng, and Jiming Chen. PointAD: Comprehending 3D anomalies from points and pixels for zero-shot 3D anomaly detection. In *NeurIPS*, volume 37, 2024b. doi: 10.52202/079017-2695.

Zhen Zhou, Le Wang, Ni Fang, Zhen Wang, Lei Qiu, and Shuo Zhang. R3D-AD: Reconstruction via diffusion for 3D anomaly detection. In *ECCV*, pp. 91–107, 2024c.

A Per-Category Comparison

Table 4: Per-category object-/point-level AUROC (O/P, %) on Anomaly-ShapeNet. C and U follow Table 1; category-specific references may be learned or nonparametric. UniSDF per-category entries use the original *raw Top-500* diagnostic readout and therefore differ in O-AUROC from the registration- and pooling-matched contrast-tail results in the main paper. Best results within C and U are bold. An em dash denotes an unreported value. Our overall values are arithmetic means of the displayed categories; the ISMP object-level mean is omitted (unavailable). External results follow the corresponding papers Zheng et al. (2025); Liang et al. (2025); Ye et al. (2025); Zha et al. (2026); Cheng et al. (2026; 2025); Kim et al. (2026); Liang et al. (2026).

Type	Method	Reference	ashtray0	bag0	bottle0	bottle1	bottle3	bowl0	bowl1	bowl2	bowl3	bowl4
C	PASDF	ICCV'25	**100.0**/91.9	**99.5**/95.8	**100.0**/95.1	**100.0/92.6**	**100.0/94.8**	**100.0**/96.3	**94.8**/90.0	**100.0**/81.6	**100.0**/93.9	93.3/86.5
C	ISMP	AAAI'25	–/60.3	–/74.7	–/77.0	–/56.8	–/77.5	–/85.1	–/54.6	–/73.6	–/77.3	–/74.0
C	PO3AD	CVPR'25	**100.0/96.2**	83.3/94.9	90.0/91.2	93.3/84.4	92.6/88.0	92.2/97.8	82.9/91.4	83.3/91.8	88.1/93.5	**98.1**/96.7
C	CASL	AAAI'26	94.3/88.7	94.8/**96.8**	95.7/85.3	95.4/86.2	**100.0**/91.4	**100.0**/95.8	93.3/**97.5**	99.3/**98.2**	99.6/**99.8**	94.4/**99.6**
C	Simple3D	AAAI'26	99.5/92.0	88.1/95.4	97.6/**97.4**	95.1/72.8	**100.0**/83.8	**100.0/98.8**	83.0/95.1	71.1/93.3	91.1/99.3	73.0/92.9
U	MC3D-AD	IJCAI'25	96.2/80.1	80.5/81.5	79.5/89.5	70.9/88.3	75.6/90.1	93.0/82.7	**97.8**/53.4	71.9/60.7	88.5/78.3	91.1/65.4
U	SeDiR	CVPR'26	97.6/76.4	88.6/89.8	92.9/91.1	91.6/89.8	98.4/**94.5**	95.9/86.6	95.6/70.0	92.6/**79.5**	96.7/86.9	**98.5**/80.6
U	MFF-M3AD	NN'26	89.0/79.1	80.0/84.8	85.2/92.2	76.8/87.5	94.0/93.0	99.2/83.3	92.2/56.1	86.7/73.4	93.3/82.8	84.8/73.1
U	UniSDF (ours)	Ours	**100.0**/72.6	89.5/95.7	**100.0/95.0**	**100.0**/90.3	96.2/91.5	**100.0**/95.5	75.6/65.9	93.3/67.8	84.8/85.0	**98.5**/80.8
U	UniNorm (ours)	Ours	**100.0/90.3**	**98.6/95.8**	**100.0**/94.9	**100.0/91.9**	**100.0**/91.5	**100.0/96.9**	94.8/**92.0**	**100.0**/74.1	**100.0/92.6**	98.2/**91.7**

Type	Method	Reference	bowl5	bucket0	bucket1	cap0	cap3	cap4	cap5	cup0	cup1	eraser0
C	PASDF	ICCV'25	**91.2**/90.9	96.8/**87.5**	77.5/82.4	85.2/94.8	64.9/86.1	64.6/89.4	**85.3**/92.0	97.1/94.8	**85.7**/88.4	95.2/94.5
C	ISMP	AAAI'25	–/53.4	–/52.4	–/67.2	–/86.5	–/73.4	–/75.3	–/67.8	–/86.9	–/60.0	–/70.6
C	PO3AD	CVPR'25	84.9/94.1	85.3/75.5	78.7/89.9	87.7/95.7	85.9/94.8	79.2/94.0	67.0/86.4	87.1/90.9	83.3/93.2	99.5/**97.4**
C	CASL	AAAI'26	88.8/97.4	**99.0**/79.6	**91.4/92.7**	**94.8**/97.4	82.1/**97.1**	77.2/94.9	55.4/66.9	99.0/**98.0**	64.8/79.7	99.5/87.6
C	Simple3D	AAAI'26	86.3/**97.9**	95.9/72.5	79.0/92.1	85.2/**98.8**	**86.7**/96.4	**91.2/97.9**	80.4/**96.4**	**100.0**/97.9	82.4/**93.7**	**100.0**/97.0
U	MC3D-AD	IJCAI'25	75.4/55.2	89.8/79.0	78.4/89.5	79.3/86.7	70.1/92.2	83.5/88.0	76.1/88.5	74.3/82.9	95.2/71.5	77.6/85.6
U	SeDiR	CVPR'26	97.9/64.6	91.1/71.7	**93.3**/88.4	**98.5**/90.0	**98.9/97.2**	**99.6/94.0**	87.4/**95.0**	**99.5**/87.1	**100.0**/84.2	66.7/74.8
U	MFF-M3AD	NN'26	98.9/56.8	84.1/72.1	90.8/86.3	89.3/88.1	81.8/90.8	97.2/91.0	**95.4**/91.9	97.6/84.9	97.6/74.8	74.3/82.3
U	UniSDF (ours)	Ours	83.9/71.1	98.7/80.8	83.8/81.5	**98.5**/92.4	61.4/76.9	45.3/77.8	77.9/74.6	93.3/92.1	78.1/76.5	94.8/93.1
U	UniNorm (ours)	Ours	**100.0/91.8**	**99.4/83.0**	86.0/**90.3**	98.2/**94.1**	80.4/87.9	50.9/88.5	84.9/91.6	93.8/**94.5**	86.2/**89.4**	**99.1/95.0**

Type	Method	Reference	headset0	headset1	helmet0	helmet1	helmet2	helmet3	jar0	microphone0	shelf0	tap0
C	PASDF	ICCV'25	**100.0**/86.3	79.5/89.1	**81.2**/81.6	93.8/64.6	76.5/80.9	84.6/95.8	**100.0**/95.9	**100.0**/95.1	71.3/86.5	**88.2/88.4**
C	ISMP	AAAI'25	–/58.0	–/70.2	–/68.3	–/62.2	–/84.4	–/72.2	–/82.3	–/66.1	–/68.7	–/52.2
C	PO3AD	CVPR'25	80.8/82.3	92.3/90.7	76.2/87.8	**96.1/94.8**	86.9/93.2	75.4/84.6	86.6/87.1	77.6/81.0	57.3/66.3	74.5/78.3
C	CASL	AAAI'26	83.6/77.0	**96.7**/90.9	78.3/**92.3**	89.0/88.2	**88.1**/92.9	**86.1/98.0**	99.5/97.4	98.1/83.8	**83.2/92.8**	76.1/65.2
C	Simple3D	AAAI'26	98.2/**93.6**	95.7/**95.8**	69.0/90.2	71.9/90.2	75.4/**94.7**	65.5/92.8	90.5/**97.8**	**100.0/96.4**	74.8/90.1	70.3/85.7
U	MC3D-AD	IJCAI'25	86.2/66.6	88.6/70.9	67.2/74.4	**100.0**/58.2	60.9/81.7	97.9/60.2	97.1/84.8	91.9/85.4	84.1/65.8	94.5/52.4
U	SeDiR	CVPR'26	84.9/70.8	89.0/68.7	**96.5/81.0**	**100.0**/59.6	87.0/**88.5**	**100.0**/70.1	**100.0**/91.1	90.5/89.1	88.7/67.9	92.7/59.7
U	MFF-M3AD	NN'26	85.8/67.1	89.5/68.9	92.5/76.1	–/54.7	**100.0**/82.3	73.3/68.2	99.0/87.0	94.8/87.6	**89.3**/62.8	**95.5**/79.2
U	UniSDF (ours)	Ours	89.3/68.3	77.6/79.8	85.5/76.2	97.6/62.6	62.9/73.0	87.9/89.3	**100.0**/94.3	91.4/90.1	68.4/77.4	67.6/77.8
U	UniNorm (ours)	Ours	**100.0/91.1**	**90.5/84.0**	88.7/78.1	**100.0/94.1**	72.2/83.1	97.9/**95.5**	**100.0/94.8**	**96.2/91.6**	71.6/**82.1**	81.8/**93.2**

Type	Method	Reference	tap1	vase0	vase1	vase2	vase3	vase4	vase5	vase7	vase8	vase9	Overall mean
C	PASDF	ICCV'25	**79.3/90.2**	**100.0**/94.4	**92.9**/79.7	**100.0**/95.6	80.6/86.8	**91.2**/89.9	**100.0**/91.5	**100.0**/95.9	92.4/90.9	**83.6**/86.3	**90.7**/89.6
C	ISMP	AAAI'25	–/55.2	–/66.1	–/84.3	–/73.3	–/76.2	–/54.5	–/47.2	–/70.1	–/85.1	–/61.5	–/69.1
C	PO3AD	CVPR'25	68.1/69.2	85.8/**95.5**	74.2/88.2	95.2/97.8	82.1/88.4	67.5/90.2	85.2/93.7	96.6/98.2	73.9/95.0	83.0/**95.2**	83.9/89.9
C	CASL	AAAI'26	56.3/60.2	83.8/89.4	89.0/**97.2**	98.1/97.1	**93.0/93.1**	82.1/91.3	71.4/73.2	**100.0/99.7**	**93.3/98.9**	81.8/88.6	88.7/89.9
C	Simple3D	AAAI'26	59.6/81.1	95.4/93.4	82.4/80.7	87.1/**98.6**	81.2/91.0	86.4/**98.5**	96.2/**96.6**	89.5/99.0	85.5/97.5	81.5/91.3	86.0/**92.9**
U	MC3D-AD	IJCAI'25	97.0/53.7	79.5/87.0	85.7/72.6	92.9/82.2	76.1/79.7	87.6/78.6	76.1/64.7	93.8/61.6	67.0/88.5	73.6/77.6	83.6/75.9
U	SeDiR	CVPR'26	89.3/63.0	88.8/84.6	97.1/72.2	97.6/85.0	78.8/83.0	**92.4/88.4**	94.8/76.6	**100.0**/77.7	87.3/90.0	**94.2**/83.5	**93.3**/81.1
U	MFF-M3AD	NN'26	**98.1**/63.8	92.9/85.5	88.1/64.0	85.2/78.2	80.6/84.8	88.2/79.2	**100.0**/56.4	64.4/73.5	75.6/89.7	83.9/74.6	89.9/77.1
U	UniSDF (ours)	Ours	73.7/77.2	**100.0**/83.7	81.4/79.1	**100.0**/94.0	80.9/83.9	86.1/68.3	91.4/72.9	**100.0**/89.0	75.2/85.5	90.3/75.1	86.5/81.4
U	UniNorm (ours)	Ours	95.9/**94.5**	**100.0/95.2**	**100.0/83.2**	**100.0/95.6**	**82.7/88.9**	79.1/**88.4**	**100.0/92.3**	**100.0/95.4**	**89.4/93.1**	87.6/**89.1**	92.6/**90.5**

Table 5: Per-category object-/point-level AUROC (O/P, %) on the 12-category Real3D-AD benchmark. Notation follows Table 4; citations are given with each external method.

Type	Method	Reference	airplane	car	candybar	chicken	diamond	duck	fish	gemstone	seahorse	shell	starfish	toffees	Overall mean
C	PASDF Zheng et al. (2025)	ICCV'25	62.8/77.7	95.9/80.2	78.8/54.6	73.9/76.8	89.4/69.9	65.8/78.2	**98.9**/83.7	63.4/65.4	**100.0**/88.7	**85.0**/64.8	61.7/70.3	86.6/83.8	80.2/74.5
C	ISMP Liang et al. (2025)	AAAI'25	**85.8**/75.3	73.1/83.6	**85.2**/90.7	71.4/79.8	94.8/92.6	71.2/87.6	94.5/88.6	46.8/85.7	72.9/81.3	62.3/83.9	66.0/64.1	84.2/89.5	75.7/83.6
C	PO3AD Ye et al. (2025)	CVPR'25	80.4/71.5	65.4/57.4	78.5/76.3	68.6/56.3	80.1/61.4	82.0/64.4	85.9/87.7	69.3/53.0	75.6/58.5	80.0/54.2	75.8/65.9	77.1/72.9	76.6/65.0
C	CASL Zha et al. (2026)	AAAI'26	80.8/84.2	79.9/90.5	84.8/93.2	65.7/71.3	97.6/98.8	**83.6**/89.5	93.5/93.5	**76.9**/91.6	64.3/81.4	79.1/**87.3**	**89.3**/83.9	**92.4**/93.7	**82.3**/88.2
C	Simple3D Cheng et al. (2026)	AAAI'26	76.5/**88.1**	**98.1/99.2**	65.1/**96.2**	**82.6/86.1**	**100.0/99.0**	77.8/**96.6**	91.2/**96.2**	70.4/**97.3**	93.0/**94.2**	51.4/71.6	69.6/**85.8**	88.8/**96.8**	80.4/**92.3**
U	MC3D-AD Cheng et al. (2025)	IJCAI'25	85.0/62.8	74.9/81.9	83.0/91.0	71.5/64.0	95.5/94.2	83.1/82.2	86.5/93.2	56.0/45.8	71.6/65.9	80.3/77.8	76.6/69.0	73.8/93.4	78.2/76.8
U	SeDiR Kim et al. (2026)	CVPR'26	86.0/66.6	78.3/**90.8**	81.9/95.1	72.9/63.2	94.8/**97.7**	**86.2/83.4**	93.8/**96.7**	62.7/54.7	67.4/73.9	77.9/79.8	85.4/70.9	84.5/**94.0**	81.0/80.6
U	MFF-M3AD Liang et al. (2026)	NN'26	**90.2**/61.5	77.0/88.9	85.1/**96.4**	**74.5**/60.0	**98.4**/96.0	84.9/81.0	94.3/95.9	59.8/54.1	76.9/72.0	74.5/78.9	80.0/68.1	78.0/92.4	81.1/78.9
U	UniSDF (ours)	Ours	60.4/78.1	80.8/74.2	82.0/55.0	74.2/76.1	90.2/69.2	67.8/78.4	**100.0**/82.7	53.8/65.6	**100.0**/87.1	72.2/65.7	88.1/86.7	84.8/83.8	79.5/75.2
U	UniNorm (ours)	Ours	64.6/**78.2**	**86.7**/75.0	**96.1**/84.6	71.2/**86.5**	95.7/86.0	74.3/78.0	98.7/84.9	**72.0/81.4**	99.6/**88.8**	**85.7/87.0**	**90.6/87.9**	**86.5**/83.5	**85.1/83.5**

B Failure Cases and Defect Families

Table 6 reports results by defect family. NULB improves object AUROC in all six named families, with the largest gains over UniSDF on scratches and holes, but broken-defect P-AUROC drops 9.34 points relative to UniNorm-Abs. A failed Chamfer alignment would trigger re-registration or reacquisition in deployment. No test sample was excluded from the reported metrics. Table 11(a) reports the flag rates in the separate full-pipeline stress study, including its clean condition.

Table 6: Results by defect type on Anomaly-ShapeNet (%). C is the number of eligible categories. Every official defect folder is included; bold denotes the best of the three residuals.

		UniSDF (auxiliary)		UniNorm-Abs		UniNorm (NULB)	
Defect	C	O	P	O	P	O	P
Bending	7	92.38	85.04	94.29	**88.83**	**96.19**	87.58
Broken	18	60.74	60.43	64.63	**62.17**	**67.22**	52.83
Bulge	40	93.38	84.43	96.26	92.33	**96.62**	**92.55**
Concavity	40	90.07	80.04	95.21	91.25	**95.88**	**91.57**
Hole	19	46.32	48.27	66.67	56.15	**68.07**	**60.41**
Scratch	19	69.12	69.74	87.63	85.39	**90.96**	**86.79**
Other (`anomaly`)	13	**65.13**	52.76	58.97	60.68	57.69	**64.15**

C Template and Reference Construction

C.1 Template provenance (train-split audit)

The audit counts released source files (four per Anomaly-ShapeNet class and one per Real3D-AD class); main scoring uses one reference per class, or 52 total. Table 8 uses UniNorm-Abs to isolate reference construction without NULB on Anomaly-ShapeNet. Independently, a provenance audit verifies that released templates are train-derived on both benchmarks (Table 7). Main results do not rely on test-time CAD privileges beyond the official train templates. Quantitatively, replacing the official template by a single train template changes Real3D-AD O-AUROC by only 0.24 points at the main readout (Appendix C).

Table 7: Template provenance audit. All sources lie in the official train splits. Relation compares released nominal geometry against the train template file.

Dataset	Templates	Source split	Relation
Anomaly-ShapeNet	160 (40 cls $\times$ 4)	train	cleaned near-ident. ($p95{=}0$)
Real3D-AD	12	train	6 identical, 5 near-ident., 1 derived

C.2 Reference source study

Table 8 uses UniNorm-Abs to isolate reference construction without NULB. On Anomaly-ShapeNet ($K{=}500$ diagnostic) a single normal training cloud is within 0.01 O-AUROC of the mesh template. On Real3D-AD at the main readout ($K{=}100$, contrast-$Q_{0.95}$), a single train template is within **0.24 O-AUROC / 0.13 O-AP** of the official mesh, and point-level AUROC is slightly higher. Object-selected atlases remain competitive, whereas medoid references are weaker because flexible match-

ing can absorb true anomalies. Together with the provenance audit (Table 7), this shows that main results do not depend on privileged external CAD.

Table 8: Reference-source study with UniNorm-Abs (%). AS uses the K=500 diagnostic grid; R3D uses the main readout (K=100, contrast-$Q_{0.95}$). Bold marks the best value in each column, including ties at the displayed precision.

	AS (40)			R3D (12)		
Reference	O	AP	P	O	AP	P
Official mesh / released template	**91.62**	**93.22**	**90.30**	84.70	81.76	83.48
Single train cloud / template	**91.62**	**93.22**	89.20	84.46	81.63	83.77
Object-selected atlas	90.96	92.78	89.83	**85.62**	83.04	**83.91**
Pointwise atlas	90.52	92.11	85.59	85.49	**84.30**	79.72
Medoid train	84.65	87.77	82.97	79.14	77.49	76.40

D Additional Analyses

Algorithm 1 UniNorm Inference

Require: Cloud X, category c, template T_c, $K = 100$, $q = 0.95$
Ensure: Point map $\mathbf{r}$, object score $A(X, c)$

Estimate $G_{X,c}$ by RANSAC followed by ICP on (X, T_c)
$\widetilde{X} \leftarrow G_{X,c}(X)$
for each $\widetilde{x}_i \in \widetilde{X}$ **do**
  $j \leftarrow \arg\min_\ell \|\widetilde{x}_i - t_\ell\|_2$
  $\delta_i \leftarrow \widetilde{x}_i - t_j$; $z_i \leftarrow \langle \delta_i, n_j \rangle$; $a_i \leftarrow |z_i|$; $\tau_i \leftarrow \|\delta_i - z_i n_j\|_2$
  $s_j \leftarrow \min(0.5u_j, 0.95)$; $r_i \leftarrow [\sqrt{1 - s_j^2}\, a_i - s_j \tau_i]_+$
end for
$A(X, c) \leftarrow A_{\mathrm{ct}}(\mathbf{r})$
return $\mathbf{r}, A(X, c)$

Table 9: Object-level readout comparison (%). Each cell is O-AUROC/P-AUROC. Learned and fixed variants share the UniNorm-Abs point map; the auxiliary class-conditioned UniSDF row uses raw Top-500 pooling, not the matched $K = 100$, contrast-Q_{95} protocol of Table 2.

Variant	ShapeNet O/P	Real3D O/P
UniSDF (auxiliary baseline)	86.52/81.35	79.53/75.22
Pseudo-selected fixed	90.35/90.30	78.98/83.48
Learned geometry	89.12/90.30	79.44/83.48
Learned SDF + geometry	82.15/90.30	77.77/83.48
UniNorm-MS	89.45/90.30	81.05/83.48
UniNorm-Abs	**91.93/90.30**	**84.70/83.48**

D.1 Residual and Aggregation Grid

Table 10 reports the residual $\times$ object-aggregation grid at $K = 100$ using the same registered inputs as Table 2. Abs. normal with contrast-Q_{95} reproduces Table 2; the Abs–SDF gap persists under identical pooling.

Table 10: Residual × object-aggregation O-AUROC (%) at $K = 100$, same matched protocol as Table 2 (Abs. normal + Contrast-Q_{95} reproduces Table 2). Bold marks the highest value in each readout column.

	Anomaly-ShapeNet			Real3D-AD		
Residual	Raw	C-Q_{95}	C-Q_{90}	Raw	C-Q_{95}	C-Q_{90}
UniSDF	86.26	85.10	85.30	80.37	81.54	81.34
Cat. SDF bank	90.73	91.43	91.19	81.31	81.94	81.91
Euclidean NN	89.77	89.96	90.11	78.97	75.87	77.47
UniNorm-Abs	**91.37**	**91.93**	**92.26**	**83.83**	**84.70**	**84.79**

D.2 Robustness Analysis

Table 11 evaluates robustness along acquisition, registration, and parameter axes. Full-pipeline tests perturb template density, coordinate noise, and surface coverage before registration; separate sweeps vary β, K_n, and K.

Selection of β. The recorded exploratory sweep evaluated $\beta \in \{0, 0.25, 0.5, 1.0\}$ using labeled Anomaly-ShapeNet test metrics, rather than a separate validation split. We selected 0.5 after inspecting both O- and P-AUROC: it had the highest P-AUROC among these settings while improving O-AUROC over $\beta = 0$. Although $\beta = 1$ had higher O-AUROC, its P-AUROC fell below the uncorrected score. Anomaly-ShapeNet is therefore a development benchmark for this NULB coefficient; Real3D-AD used the coefficient unchanged. Anomaly labels were not used to construct templates or point scores.

Table 11(a) reports the full-pipeline perturbation study, which uses UniNorm-Abs and is applied before registration. Retaining 25% of template points preserves 91.51/91.15 O/P-AUROC, whereas coordinate noise mainly harms localization. Removing 75% of a contiguous reference region triggers the Chamfer flag for every scan and collapses detection, exposing the shared alignment and coverage boundary. Accordingly, a protocol-fixed Chamfer failure should trigger re-registration or reacquisition in deployment; this operational abstention is not included in benchmark AUROC.

Table 11(b) changes only K_n on the fixed 40-class Anomaly-ShapeNet cache. Across $K_n \in \{10, 20, 30, 50, 100\}$, O/P-AUROC varies by at most 0.19/0.07 points, confirming local parameter stability but not removing the stronger dependence on registration and coverage.

Table 11: Robustness analysis (%). (a) Full-pipeline UniNorm-Abs template stress on eight Anomaly-ShapeNet classes. CD flag (%) is the percentage of scans whose Chamfer residual exceeds the fixed registration threshold $\epsilon = 1.8$. (b) NULB neighborhood sensitivity on all 40 classes; all other settings are fixed.

(a) Registration, density, and coverage stress

Condition	O-AUROC	P-AUROC	CD flag
Clean	94.26	91.29	0.34
25% template points	91.51	91.15	1.35
Noise $\sigma = 0.005$	92.46	73.24	0.00
75% partial removal	48.79	78.56	100.00

(b) Normal-neighborhood size

K_n	O-AUROC	P-AUROC
10	92.45	90.47
20	92.58	90.51
30	92.60	90.53
50	92.57	**90.54**
100	**92.64**	90.53

E Qualitative Results

Figure 4 (main text) shows the eight-example comparison: both SDF variants often retain broad background activation, whereas UniNorm concentrates evidence around the annotations. Figure 8

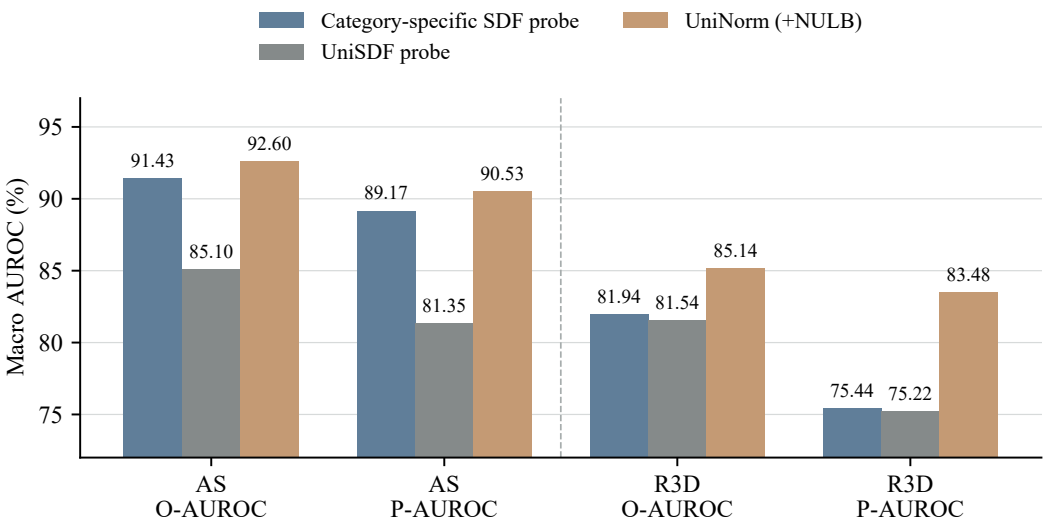

Figure 5: Residual comparison using the results in Table 2. All methods use the same registered test inputs and contrast-tail object pooling ($K = 100$, $q = 0.95$). The category-specific SDF is a diagnostic PASDF-style probe, not the published PASDF baseline. Representations differ across the probes and UniNorm. AS: Anomaly-ShapeNet; R3D: Real3D-AD.

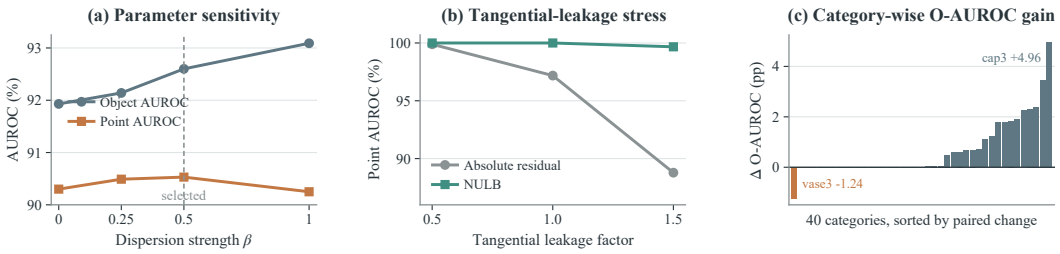

Figure 6: Complementary evidence for NULB. (a) Labeled Anomaly-ShapeNet development sweep for β; $\beta = 0.5$ was frozen for Real3D-AD. (b) Point-level AUROC under increasing tangential leakage in a controlled local simulation. (c) Paired object-level changes across the 40 Anomaly-ShapeNet categories.

provides an enlarged view. Independent 5th–99.5th percentile heat-map normalization is for display only.

F Efficiency Analysis

SDF training/inference: NVIDIA RTX 3090; registration and scoring: Intel Xeon Gold 6133 @ 2.50 GHz.

Table 12 separates neural state from measurement state. The 52 templates contain 1.04M points and require 27.70 MiB for float32 points, normals, and uncertainty. On 80 cached scans, UniNorm-Abs and NULB require 18.71 and 20.89 ms/scan, respectively. Registration requires 509.5 ms/scan and therefore remains the deployment bottleneck; real-time use needs faster or sensor-initialized alignment.

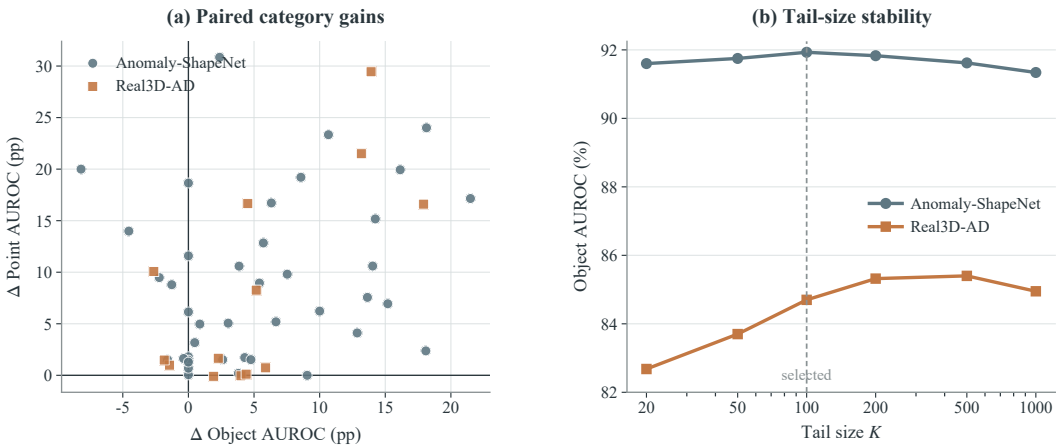

Figure 7: Residual and object-level readout analyses. (a) Paired category gains of UniNorm-Abs over the class-conditioned UniSDF baseline. (b) O-AUROC under a shared Top-K sweep; $K = 100$ is fixed for both datasets rather than optimized per category. NULB is analyzed separately in Fig. 6.

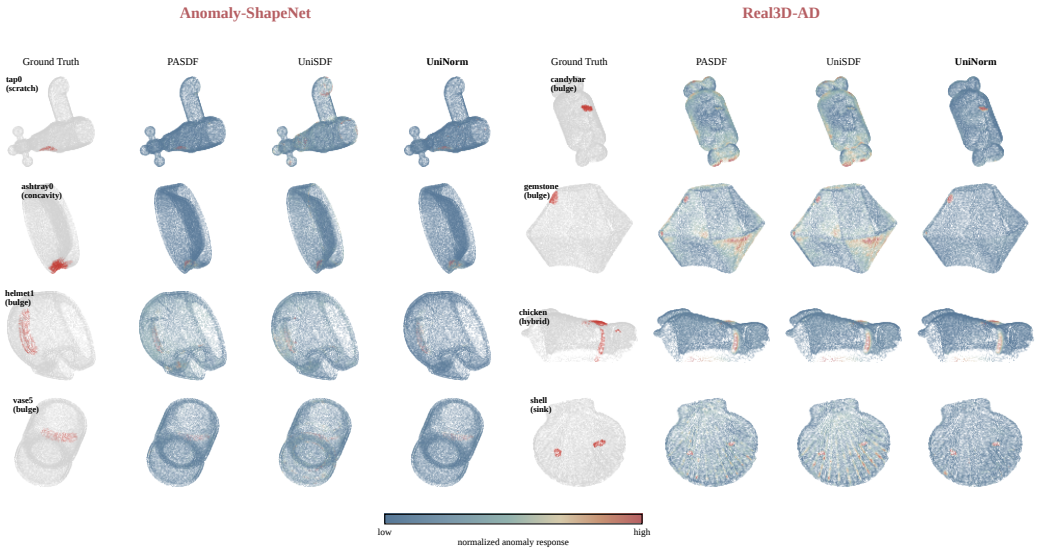

Figure 8: Enlarged version of Fig. 4: eight examples across Anomaly-ShapeNet and Real3D-AD. Each group shows ground truth, the PASDF-style SDF probe, UniSDF, and UniNorm. Each score map uses an independent 5th–99.5th percentile color scale for display only.

Table 12: State and latency. Cached-score means use 80 fixed scans and three repeats on CPU; registration is from the separate full-pipeline profile.

Resource	Value
UniSDF auxiliary parameters	3.55M
Category-specific PASDF-style SDF bank	15.87M / 4.76M
Learned scoring parameters	0
Final template points (52 classes)	1.04M
T, N, U array storage	27.70 MiB
Offline U construction (40 classes)	2.59 s
Registration / scan	509.5 ms
Cached UniNorm-Abs score / scan	18.71 ms
Cached NULB score / scan	20.89 ms
NULB scoring increment	2.18 ms (11.7%)