FROM VARIATION TO VIOLATION: LEARNING 3D ANOMALY SHAPES FROM NORMAL SURFACES

Anonymous authors
Paper under double-blind review

ABSTRACT

Existing 3D point cloud anomaly synthesis methods largely rely on hand-crafted deformation operators or predefined anomaly morphologies, limiting the diversity of synthesized anomalies. We observe that local changes such as bulges and depressions can also occur naturally across normal instances, and may become anomalous when they become stronger than what is normally observed. Based on this observation, we propose NOVA, a variation-driven framework for 3D anomaly synthesis and detection. NOVA uses local surface differences among aligned normal instances as deformation patterns and strengthens them beyond normal variation to create pseudo-anomalies. For detection, a normal-geometry branch first learns how much and in which directions normal surfaces vary, and is then frozen. Using its geometric features and reference-matching statistics, an anomaly-scoring branch learns from paired normal and deformed samples to assign higher scores to the same points after deformation. The resulting scores support anomaly detection and localization. Experiments confirm more diverse synthesized anomaly patterns. The normal-geometry branch alone surpasses the best published O-AUROC and P-AUROC results on both datasets using only normal training data. Supervision from 32 pseudo-anomalies per category further improves performance, raising P-AUROC from 92.68% to 95.99% on Real3D-AD and from 96.94% to 97.16% on Anomaly-ShapeNet. The resulting framework achieves O-AUROC scores of 91.28% and 96.37%, respectively. Code: `https://anonymous.4open.science/r/NOVA-7C92/`.

1 INTRODUCTION

3D anomaly detection (3D-AD) aims to identify anomalous objects and localize geometric irregularities on their surfaces. In industrial inspection, anomalous samples vary widely and are costly to collect, so training often relies on a small number of normal objects. Existing methods learn normal geometry through feature matching or reconstruction and detect deviations at test time (Liu et al., 2023a; Bergmann & Sattlegger, 2023; Wu et al., 2026b). The core difficulty is to separate a geometric anomaly from the shape and acquisition variation already present among normal objects.

Limitation of existing synthesis. Recent studies provide supervision by synthesizing pseudo-anomalies through patch-wise geometric perturbations (Zhou et al., 2024), surface-normal-guided point displacements (Ye et al., 2025), multiple anomaly types and scales (Hoang et al., 2026a; Lv et al., 2026; Kang et al., 2026), physics-inspired deformation functions (Ning et al., 2026), and modular parameterized geometric deformations (Balapour & Hach, 2026). These methods adapt the generated anomalies to local geometry. For displacement-based synthesis, however, predefined deformation functions largely determine how the local surface changes—for example, where it rises or dips and how the relative displacement varies across the region. Such designs therefore require explicit deformation specification and restrict synthesized anomalies to predefined deformation families. More importantly, they do not directly use variations among normal instances to determine how local surfaces naturally change and how strong such changes can normally be.

Motivation. For local geometric anomalies, normal and abnormal surfaces can differ mainly in the degree of a surface change: slight surface undulations may be normal, whereas an excessive local bulge or depression may indicate an anomaly. Local shape differences already present among aligned normal objects therefore offer a natural starting point for synthesis. NOVA draws deforma-

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.

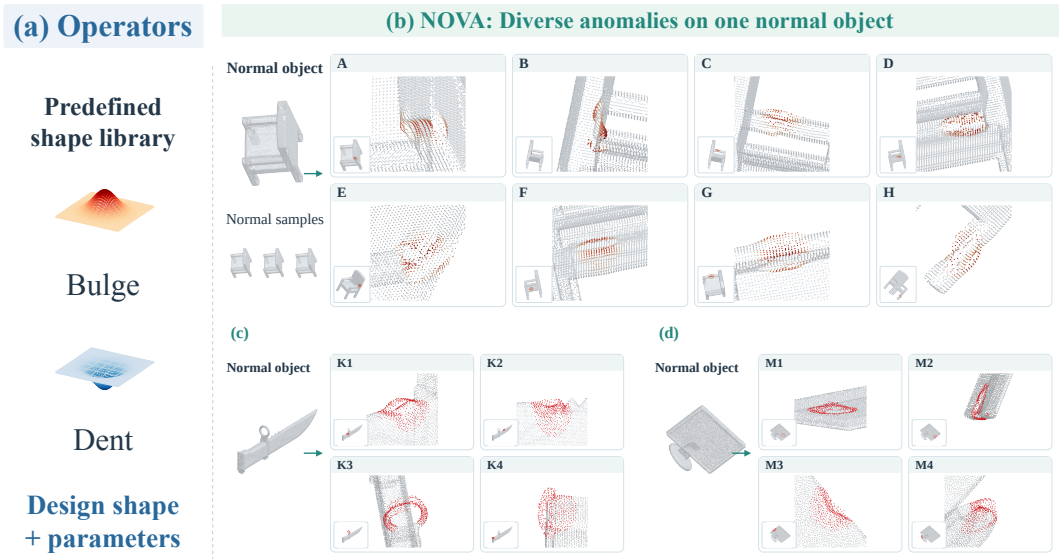


Figure 1: **Conceptual comparison of predefined anomaly synthesis and NOVA.** (a) Operator-designed shapes and parameters. (b–d) NOVA examples on a chair, knife, and monitor, with one fixed normal object per group. Insets locate each deformation. Warm colors in (b) show displacement magnitude on a shared scale; red in (c,d) marks displaced points.

tion patterns from these differences and strengthens them locally, using the amplitudes observed in normal samples as a reference, to generate pseudo-anomalies for training. Normal data thus provide both a source of deformation shapes and a reference scale for constructing changes beyond the observed normal variation (Figure 1).

From variation modeling to anomaly detection. The same view also motivates separating normal modeling from synthetic anomaly supervision. Normal samples first reveal how much and in which directions local geometry can vary. Since our synthesis preserves point correspondence, paired normal and deformed samples then provide direct supervision on how anomaly scores should increase when these variations are strengthened. This motivates learning geometric features from normal samples first, then using paired synthetic anomalies to learn how these features should translate into anomaly scores.

We introduce **NOVA** (**NO**rmal-**VA**riation-driven anomaly synthesis), which uses variation among normal surfaces to guide both anomaly synthesis and detection. For synthesis, NOVA extracts local deformation patterns from differences between aligned normal objects and strengthens them relative to normal variation. For detection, a frozen normal-geometry branch provides geometric features and reference-matching statistics to an anomaly-scoring branch trained on paired normal and deformed samples. At inference, this branch produces point scores for localization, and their maximum gives the object-level anomaly score. Only the query object and normal references are required. Experiments show that NOVA produces more diverse anomaly patterns than predefined synthesis methods. Notably, the normal-geometry branch alone, trained without synthetic anomaly supervision, already surpasses the best published object- and point-level AUROC results on both benchmarks. Adding synthetic supervision further improves point-level detection, leading to state-of-the-art detection and localization performance. Our contributions are as follows.

- **Variation-driven pseudo-anomaly synthesis.** NOVA draws local deformation patterns from surface differences between aligned normal objects, generating pseudo-anomalies of many shapes from anomaly-free training data without designing a separate shape for each anomaly. NOVA can generate at most eight distinct anomaly morphologies on a single normal object without designing a separate deformation operator for each morphology.
- **Paired anomaly detection.** To learn from synthetic anomalies while preserving the geometry learned from normal samples, we train an anomaly-scoring branch on frozen ge-

ometric features. Pointwise comparisons between normal samples and their deformed counterparts teach the branch to assign higher scores to introduced deformations. With 32 pseudo-anomalies per category, P-AUROC rises from 92.68% to 95.99% on Real3D-AD and from 96.94% to 97.16% on Anomaly-ShapeNet compared with normal-only detection. O-AUROC reaches 91.28% and 96.37%, respectively.

2 Related Work

3D point-cloud anomaly detection. Point-cloud anomaly detection includes feature comparison, reconstruction, and discrimination between normal and synthesized samples. Feature-based methods compare geometric descriptors or learned representations using normal-reference matching (Horwitz & Hoshen, 2023; Liu et al., 2023a; Yu et al., 2025) or student–teacher disagreement (Bergmann & Sattlegger, 2023). Further work addresses internal structure, subtle surface anomalies, and multi-category detection (Liang et al., 2025a; Cheng et al., 2026; Kim et al., 2026). Reconstruction methods recover normal geometry through diffusion or continuous surface representations and localize anomalies from reconstruction discrepancies (Zhou et al., 2024; Zheng et al., 2025; Wu et al., 2026b). Across these detection strategies, synthetic anomalies can provide training supervision.

Anomaly synthesis for detector training. Synthetic supervision uses image or feature perturbations (Li et al., 2021; Zavrtanik et al., 2021; Liu et al., 2023b) and depth-based anomaly simulation (Zavrtanik et al., 2024). Point-cloud methods similarly alter geometry or features. Geometric generators use local patch perturbations and normal-guided displacements (Zhou et al., 2024; Ye et al., 2025), combine multiple anomaly types (Hoang et al., 2026a; Kang et al., 2026; Lv et al., 2026), or adopt physics-inspired and modular parameterized deformations (Ning et al., 2026; Balapour & Hach, 2026). Feature-space methods perturb selected tokens to supply anomaly supervision (Hoang et al., 2026b;c). Although these methods adapt anomaly generation to local geometry, their geometric deformation patterns are largely specified by predefined operators or parameterized deformation families. NOVA derives both local deformation patterns and their strength reference from variations among aligned normal objects.

For detector training, synthetic samples provide reconstruction or point-offset targets (Liang et al., 2025b; A et al., 2026; Ye et al., 2025), support contrastive learning and adaptation before feature matching (Zhu et al., 2024; Cheng et al., 2025b), or supervise classification (Zha et al., 2026) and point–language prompt learning (Wang et al., 2026). Normal-reference comparison and contrastive learning on normal samples are also explored through comparative attention and Siamese modeling (Wu et al., 2026a). NOVA differs by separating normal-variation modeling from synthetic anomaly supervision. It first learns how normal surfaces vary using only normal data and keeps this model fixed. Paired normal and deformed samples are then used to train an anomaly-scoring branch to increase the scores of the same points after deformation.

3 Method

3.1 Overview

NOVA uses aligned anomaly-free point clouds for anomaly synthesis and detection (Figure 2). Inter-instance surface differences provide both the spatial patterns of local deformations and an empirical reference for their strength. The detector first learns local geometry and reference-matching statistics from normal samples. With the normal-geometry branch frozen, these features are used to train the anomaly-scoring branch through pointwise comparisons before and after deformation. At inference, the predicted scores are calibrated and spatially aggregated into a point-level anomaly map, whose maximum gives the object-level anomaly score.

3.2 Variation-Driven Pseudo-Anomaly Synthesis

NOVA separates a deformation's spatial pattern from its overall amplitude. This separation allows NOVA to preserve the extracted spatial pattern while controlling deformation strength independently. Let $X^s = \{x_i^s\}$ be an anomaly-free source object from category $\mathcal{C}$, and let Ω_c be a connected local

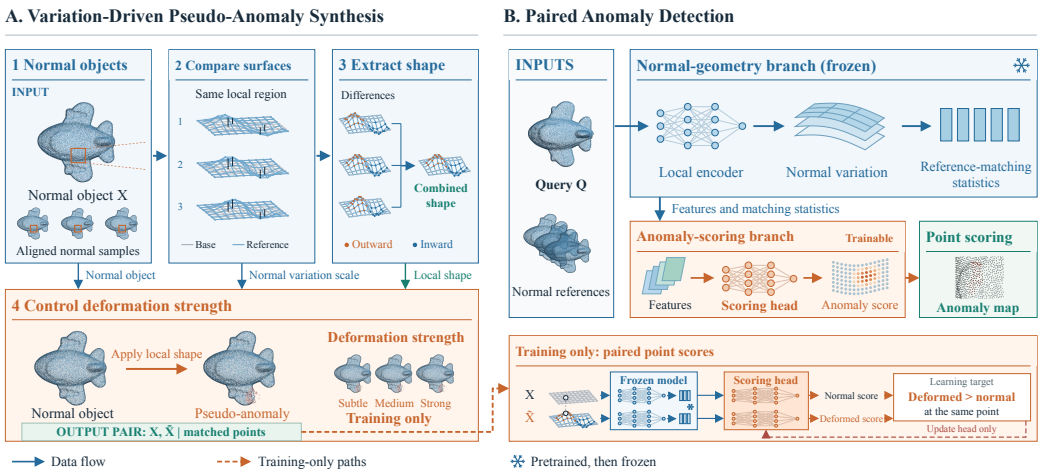


Figure 2: **Overview of NOVA.** (A) Differences between aligned normal surfaces provide local anomaly shapes, with strength controlled relative to normal variation. (B) The frozen normal-geometry branch supplies features and matching statistics; the anomaly-scoring branch learns to increase point scores after deformation and produces the anomaly map. Orange and red dashed arrows indicate training inputs and updates to the scoring branch, respectively.

region around center c. The region determines where a deformation is applied; surface differences determine its spatial pattern within that region.

Extracting deformation patterns. We compare X^s with T aligned anomaly-free target objects $\{X^t\}_{t=1}^T$. For each valid match, we measure the signed offset along the source surface normal to the matched target's tangent plane, expressed in local point-spacing units. We remove a smooth local trend from these offsets and normalize the residuals by their weighted 95th-percentile absolute amplitude, obtaining $\bar{h}_i^{(t)}$ (Appendix A).

Let w_i be a boundary-tapering weight. Among points with valid matches to all targets and $w_i > 0$, let $\rho_i^{(t)} \in [0, 1]$ be the percentile rank of $|w_i \bar{h}_i^{(t)}|$ for target t. We combine the normalized residuals and their ranks as

$$p_i = \frac{w_i}{Z} \operatorname{sign}\left(\sum_{t=1}^{T} \bar{h}_i^{(t)}\right)\left(\frac{1}{T}\sum_{t=1}^{T} \rho_i^{(t)}\right). \tag{1}$$

The pattern is set to zero outside this common valid support. Here Z is the weighted 95th percentile of the absolute numerator, floored at a small $\varepsilon > 0$ and evaluated over Ω_c, including zeroed entries. Thus, the summed residuals determine the local sign, the averaged ranks determine the relative magnitude, and w_i tapers the displacement toward the region boundary. The resulting p_i is a dimensionless signed scalar specifying relative displacement along the source surface normal.

Controlling deformation strength. To reintroduce an empirical amplitude reference after pattern normalization, we estimate $b_{\mathcal{C}}$ from surface offsets before trend removal. We first compute the weighted 95th-percentile absolute offset for each source–target region, then take the 99th percentile of these regional amplitudes within category $\mathcal{C}$. For deformation e, a positive increment Δ_e sets the amplitude coefficient to $b_{\mathcal{C}} + \Delta_e$ without changing p_i.

Constructing paired samples. We apply the scaled pattern along the source surface normals:

$$\widetilde{x}_i^s = x_i^s + \kappa_e (b_{\mathcal{C}} + \Delta_e) \ell_i^s p_i n_i^s. \tag{2}$$

Here n_i^s is the unit surface normal, ℓ_i^s is the local point spacing that converts the dimensionless displacement into coordinate units, and $\kappa_e \in \{-1, +1\}$ preserves or reverses the pattern's sign. Points outside Ω_c remain unchanged. This construction pairs the normal sample $X = X^s$ with its deformed version $\widetilde{X}$, preserving point indices. Surface matching, region construction, trend removal, and synthesis settings are detailed in Appendix A.

3.3 PAIRED ANOMALY DETECTION

The normal-geometry branch learns from normal samples and is then frozen. Its geometric features and reference-matching statistics provide the inputs to the anomaly-scoring branch, which learns from paired normal and deformed samples.

Normal-geometry branch. The same displacement magnitude can produce different geometric responses depending on the local surface and displacement direction. We account for normal displacement variation using a covariance predicted from local geometry. For each query point, we find its nearest point in each aligned anomaly-free reference from the same category. A shared local MLP encoder describes the query and reference neighborhoods. A matching network combines these descriptors and their relative geometry to predict a positive-definite covariance Σ_{ir}. The encoder and covariance predictor are trained on normal samples using a Gaussian-mixture likelihood objective, with references supplying the mixture components (Appendix B). On Real3D-AD, we further refine this branch using only normal samples before freezing it (Appendix B.5).

Let d_{ir} denote the displacement from query point i to its matched point in reference r. The Mahalanobis statistic $d_{ir}^{\top}\Sigma_{ir}^{-1}d_{ir}$ measures this displacement relative to the modeled normal variation, assigning larger values to deviations along directions with lower variability. These statistics, together with local descriptors and reference-agreement features, provide the inputs to the anomaly-scoring branch (Appendix B.2).

Anomaly-scoring branch. The anomaly-scoring branch uses an MLP head to map frozen local features and reference-matching statistics to a raw pointwise anomaly score a_i. We train this head to assign each displaced point a higher score than the same point before deformation, while keeping scores on normal samples near zero.

For a normal sample X and its deformed version $\widetilde{X}$, define $m_i = \|\widetilde{x}_i - x_i\|_2/\ell_i$, where ℓ_i is the local point spacing in the normal sample. Let X^{scan} denote a scan-perturbed anomaly-free sample, $\mathbb{E}_i$ the average over points, and $\mathrm{softplus}(v) = \log(1+e^v)$. With $\varepsilon > 0$ a numerical floor, we minimize

$$\mathcal{L}_{\mathrm{AE}} = \frac{\sum_i m_i \,\mathrm{softplus}\Big(-[a_i(\widetilde{X}) - a_i(X)]\Big)}{\max(\sum_i m_i, \varepsilon)} + \tfrac{1}{2}\mathbb{E}_i[a_i(X)^2] + \tfrac{1}{2}\mathbb{E}_i[a_i(X^{\mathrm{scan}})^2]. \tag{3}$$

The ranking term gives larger displacements greater weight and unchanged points zero weight. The two squared-score terms apply to normal samples and their scan-perturbed versions. Only the anomaly-scoring head is updated; the encoder and covariance predictor remain frozen.

3.4 POINT AND OBJECT SCORING

We calibrate the raw point scores using normal samples from each category. Let q_{95}^{A} and q_{99}^{A} denote their 95th and 99th percentiles after head training. The calibrated response is

$$E_i = \mathrm{ReLU}\left(\frac{a_i - q_{99}^{A}}{\max(q_{99}^{A} - q_{95}^{A}, \varepsilon)}\right). \tag{4}$$

Here E_i retains the positive excess above the normal 99th percentile. Calibration uses clean and scan-perturbed normal samples (Appendix B.3).

Let $\mathcal{A}$ denote two rounds of averaging over 16-point neighborhoods, each including the center point, followed by softplus. We spatially aggregate the calibrated response to obtain the point anomaly map:

$$S_i = \mathcal{A}(E)_i, \qquad S_{\mathrm{obj}}(Q) = \max_{1 \le i \le |Q|} S_i. \tag{5}$$

Here Q is the query cloud and $|Q|$ its number of points. The point anomaly map supports localization, and its maximum gives the object-level anomaly score. Inference requires only the query and its category's anomaly-free references.

4 EXPERIMENTS

We evaluate overall performance, anomaly synthesis, and paired anomaly detection.

4.1 EXPERIMENTAL SETUP

Datasets and metrics. We evaluate on all 12 Real3D-AD categories (Liu et al., 2023a) and the original 40 Anomaly-ShapeNet categories (Li et al., 2024), using four normal training objects per category. We report object- and point-level AUROC and average precision (O-AUROC, P-AUROC, O-AP, and P-AP), macro-averaged across categories.

Implementation details. No real anomalous samples are used for training or calibration; all pseudo-anomalies are synthesized from the normal training objects. Unless otherwise stated, main results use 32 pseudo-anomalies per category on both datasets and are averaged over three head-training seeds, with the normal-geometry model and synthesis pool fixed. Additional training details and registration sensitivity are provided in Appendices B.4 and D.2.

4.2 COMPARISON WITH EXISTING METHODS

Quantitative results. Table 1 compares NOVA with 20 published point-cloud anomaly detection methods, grouped by training supervision and anomaly synthesis. Even without synthetic anomaly supervision, the normal-geometry branch reaches O-/P-AUROC of 88.54%/92.68% on Real3D-AD and 96.12%/96.94% on Anomaly-ShapeNet, surpassing the best previously published O-/P-AUROC results on both datasets. With 32 pseudo-anomalies per category, NOVA further improves P-AUROC to 95.99% and 97.16%, with O-AUROC reaching 91.28% and 96.37%, respectively. These means over three head-training seeds are the highest reported values in Table 1.

Qualitative comparison. Figure 3 compares 12 selected objects with CASL (Zha et al., 2026) and PA3AD (Ning et al., 2026). NOVA has higher point AP on the eight anomalous objects and lower point FPR on the four normal objects shown with PA3AD. The CASL panels compare anomaly localization using relative score maps; the PA3AD panels also assess false positives with thresholds estimated from training-normal samples. Appendix C.3 gives the display and selection protocol.

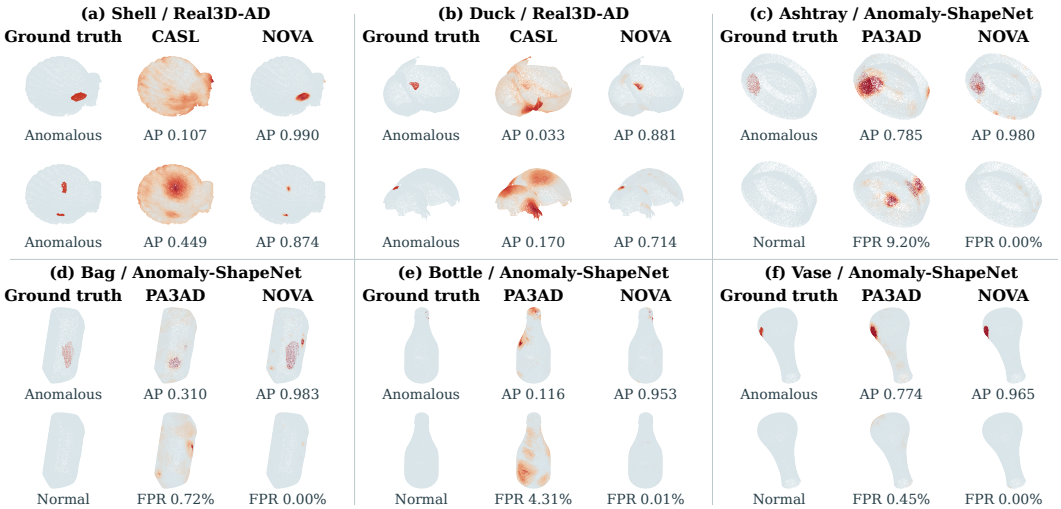


Figure 3: **Selected detection comparisons with CASL (a,b) and PA3AD (c–f).** NOVA improves point AP on anomalous objects and reduces point FPR on normal objects. Red denotes annotated anomalies.

Table 1: Published-method comparison of object detection (O) and point localization (P), grouped by anomaly synthesis. NOVA has the highest reported means in all four columns. AUROC (higher is better). The best results are shown in bold, and the second-best results are underlined. Protocol notes are provided in Appendix C.1.

Method	Venue	Anomaly synthesis	Real3D-AD		Anomaly-ShapeNet	
			O	P	O	P
Normal-only detection						
PatchCore (Roth et al., 2022)	CVPR 2022	None (normal data only)	0.594	0.620	0.562	0.577
Reg3D-AD (Liu et al., 2023a)	NeurIPS 2023	None (normal data only)	0.704	0.705	0.572	0.668
IMRNet (Li et al., 2024)	CVPR 2024	None (normal data only)	0.725	–	0.661	0.650
CPMF (Cao et al., 2024)	PR 2024	None (normal data only)	0.645	<u>0.930</u>	0.559	0.573
ISMP (Liang et al., 2025a)	AAAI 2025	None (normal data only)	0.767	0.836	–	0.691
MC3D-AD (Cheng et al., 2025a)	IJCAI 2025	None (normal data only)	0.782	0.768	0.842	0.748
PASDF (Zheng et al., 2025)	ICCV 2025	None (normal data only)	0.802	0.745	0.900	0.897
Reg2Inv (Yu et al., 2025)	NeurIPS 2025	None (normal data only)	0.780	0.878	0.861	0.882
Simple3D (Cheng et al., 2026)	AAAI 2026	None (normal data only)	0.804	0.923	0.860	0.929
SeDiR (Kim et al., 2026)	CVPR 2026	None (normal data only)	0.810	0.806	0.933	0.810
NOVA (Normal-only)	–	None (normal data only)	<u>0.885</u>	0.927	<u>0.961</u>	<u>0.969</u>
Rule-based synthesis						
Group3AD (Zhu et al., 2024)	ACM MM 2024	Local Gaussian noise	0.751	0.735	–	–
R3D-AD (Zhou et al., 2024)	ECCV 2024	Patch displacement (Patch-Gen)	0.734	0.592	0.749	–
PO3AD (Ye et al., 2025)	CVPR 2025	Normal-guided offsets (Norm-AS)	0.765	0.650	0.839	0.898
DUS-Net (Liang et al., 2025b)	ACM MM 2025	Center/patch Gaussian noise	0.795	0.847	0.797	0.712
CASL (Zha et al., 2026)	AAAI 2026	Normal-guided offsets (PO3AD)	0.823	0.882	0.887	0.899
AARD (Wu et al., 2026b)	CVPR 2026	Directional deformation (DA-Gen)	0.820	0.860	0.931	0.936
HPPF (Kang et al., 2026)	CVPR 2026	Kernel/sine offsets; region removal	0.842	0.874	0.876	0.912
PA3AD (Ning et al., 2026)	PR 2026	Physics-inspired deformation functions	0.789	0.839	0.933	0.931
PO3AD+AF3AD (Balapour & Hach, 2026)	ECCV 2026	Parametric deformation families	0.852	0.861	0.915	0.925
MAD-3D (Lv et al., 2026)	IVC 2026	Local transformations + noise queries	0.773	0.836	0.903	0.950
Normal-data-driven synthesis						
NOVA	–	Aligned normal-surface differences	**0.913**	**0.960**	**0.964**	**0.972**

4.3 Effectiveness of Anomaly Synthesis

Adding the anomaly-scoring branch trained on synthetic pairs improves all four metrics over normal-only detection on both datasets (Table 3; Section 4.4). We examine how synthesis source and quantity affect detection, and how the generated shapes vary with strength and location.

Comparison of synthesis methods. Table 2 compares anisotropic Gaussian deformation, strategies adapted from PO3AD, CASL, AF3AD, and PA3AD, and NOVA. Each supplies 16/32 pseudo-anomalies per Real3D-AD/Anomaly-ShapeNet category, with the normal model, head architecture, training budget, and scorer fixed. The Gaussian control forms a smooth protrusion or indentation on the same normal objects and regions, using NOVA's strength schedule to compare predefined shapes with normal-surface-derived shapes. Published methods retain their own region and deformation rules (Appendix C.2).

On Real3D-AD, NOVA leads O-AUROC, P-AUROC, and P-AP, while AF3AD leads O-AP. NOVA achieves P-AUROC/P-AP of 96.16%/52.49%, compared with 94.59%/45.91% for Gaussian deformation. On Anomaly-ShapeNet, CASL leads O-AUROC (96.76%) and P-AP (78.12%), compared with NOVA's 96.37% and 76.82%; Gaussian deformation leads P-AUROC. No tested strategy is best across both datasets. Different deformation patterns and strengths may match real anomalies to different degrees, but these properties vary together here, so their individual effects remain to be established.

Diversity of synthesized shapes. Figure 1 shows local anomalies that vary in location and shape on the same chair, without a separately designed operator for each shape; Appendix A.3 gives examples on other objects. To quantify shape variation, we compare 16 matched pairs of NOVA and Gaussian deformations on the same chair. After removing mean displacement and overall amplitude, spatial-frequency distributions describe their broad and fine-scale changes. Under this measure, the

Table 2: **Comparison of anomaly synthesis methods with a fixed NOVA detector and equal sample counts.** No method achieves the best performance across both datasets.

Synthesis	Venue	Real3D-AD				Anomaly-ShapeNet			
		O-AUROC	O-AP	P-AUROC	P-AP	O-AUROC	O-AP	P-AUROC	P-AP
Gaussian deformation	–	0.9015 ±0.0008	0.8883 ±0.0018	0.9459 ±0.0011	0.4591 ±0.0060	0.9660 ±0.0007	0.9805 ±0.0002	**0.9759** ±0.0003	0.7742 ±0.0016
PO3AD (Ye et al., 2025)	CVPR 2025	0.9042 ±0.0011	0.8869 ±0.0019	0.9542 ±0.0008	0.4928 ±0.0012	0.9645 ±0.0013	0.9802 ±0.0009	0.9753 ±0.0004	0.7799 ±0.0024
CASL (Zha et al., 2026)	AAAI 2026	0.8970 ±0.0025	0.8825 ±0.0011	0.9487 ±0.0021	0.4678 ±0.0107	**0.9676** ±0.0007	**0.9817** ±0.0003	0.9752 ±0.0005	**0.7812** ±0.0043
AF3AD (Balapour & Hach, 2026)	ECCV 2026	0.9131 ±0.0020	**0.9002** ±0.0013	0.9538 ±0.0023	0.4970 ±0.0116	0.9626 ±0.0014	0.9795 ±0.0007	0.9755 ±0.0001	0.7749 ±0.0033
PA3AD (Ning et al., 2026)	PR 2026	0.8954 ±0.0080	0.8786 ±0.0038	0.9577 ±0.0032	0.5026 ±0.0157	0.9648 ±0.0019	0.9804 ±0.0009	0.9732 ±0.0013	0.7737 ±0.0024
Ours	–	**0.9155** ±0.0035	0.8957 ±0.0023	**0.9616** ±0.0007	**0.5249** ±0.0055	0.9637 ±0.0018	0.9800 ±0.0009	0.9716 ±0.0012	0.7682 ±0.0017

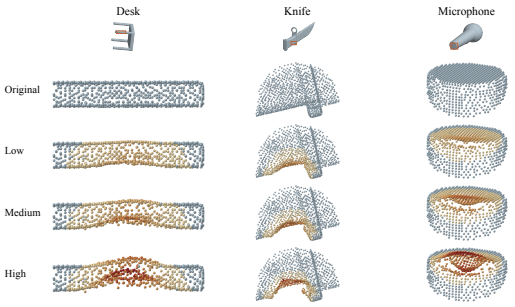


Figure 4: **Controllable anomaly strength.** Three objects are shown before deformation and at low, medium, and high strengths, with each region and deformation pattern fixed.

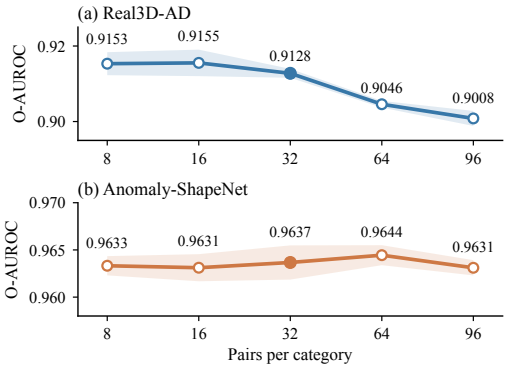


Figure 5: **Effect of synthesis quantity.** Different synthesis quantities are compared under the same detector settings. More anomalies do not consistently improve detection.

mean pairwise difference among NOVA deformations is 15.6 times that of the Gaussian control, indicating greater shape variation after controlling for amplitude.

Controllable deformation strength. Figure 4 shows local deformations of a desk, a knife, and a microphone at increasing strengths. Within each example, the object, region, and deformation pattern remain fixed; increasing only the added amplitude makes the local shape change more pronounced. This illustrates how NOVA controls deformation strength independently of the spatial pattern extracted from normal surfaces.

Number of synthetic training pairs. We examine 8, 16, 32, 64, and 96 pseudo-anomalies per category over three head-training seeds (Figure 5). Mean O-AUROC peaks at 16 pairs on Real3D-AD and 64 on Anomaly-ShapeNet; the highest point-level metrics occur at these counts as well. Increasing from 64 to 96 lowers mean O-AUROC from 90.46% to 90.08% and from 96.44% to 96.31%, respectively. All heads train from scratch for eight epochs, so larger pools entail more updates (full metrics in Appendix D.1). Additional controls show that the effect of synthesis quantity depends on the training budget, while longer training and a wider scoring head provide no consistent improvement across metrics and datasets (Appendix D.1.1).

4.4 Effectiveness of Paired Anomaly Detection

Comparison with normal-only detection. The normal-geometry branch alone already reaches O-/P-AUROC of 88.54%/92.68% on Real3D-AD and 96.12%/96.94% on Anomaly-ShapeNet, sur-

passing the best previously published O-/P-AUROC results in Table 1. Normal-only uses the calibrated geometric response of the frozen normal-geometry model (Appendix B.3). Table 3 compares Normal-only with NOVA while keeping that model fixed. NOVA uses the frozen geometric features to learn anomaly scores from 32 synthetic pairs per category. P-AUROC/P-AP rises from 92.68%/42.58% to 95.99%/52.18% on Real3D-AD and from 96.94%/73.73% to 97.16%/76.82% on Anomaly-ShapeNet. These results indicate that the synthesized anomalies provide useful supervision for learning anomaly scores from normal geometric features.

Table 3: **Detection with and without synthetic supervision.** With the normal model fixed, NOVA learns point scores from 32 pseudo-anomalies per category. It improves all four metrics over Normal-only on both datasets.

Output	Real3D-AD				Anomaly-ShapeNet			
	O-AUROC	O-AP	P-AUROC	P-AP	O-AUROC	O-AP	P-AUROC	P-AP
Normal-only	0.8854	0.8704	0.9268	0.4258	0.9612	0.9795	0.9694	0.7373
NOVA	**0.9128** ±0.0012	**0.8959** ±0.0005	**0.9599** ±0.0006	**0.5218** ±0.0074	**0.9637** ±0.0018	**0.9800** ±0.0009	**0.9716** ±0.0012	**0.7682** ±0.0017

Comparison with other detectors. Table 4 compares detectors trained on the same NOVA pseudo-anomalies, with 16 pairs per Real3D-AD category and 32 per Anomaly-ShapeNet category. NOVA achieves higher object- and point-level metrics than the evaluated baselines in this fixed-data setting. This comparison highlights the importance of how a detector uses synthetic supervision, beyond the availability of synthetic anomalies alone. By retaining a normal-geometry reference and learning from paired synthetic anomalies, NOVA captures both what normal geometry looks like and which local changes indicate anomalies.

Table 4: **Comparison of detectors trained on the same synthetic data.** Each detector retains its own training and scoring procedures. NOVA achieves the highest metrics among the reported runs.

Detector	Real3D-AD				Anomaly-ShapeNet			
	O-AUROC	O-AP	P-AUROC	P-AP	O-AUROC	O-AP	P-AUROC	P-AP
PointNet++	0.5101	0.5464	0.5061	0.0147	0.5378	0.6121	0.5389	0.0299
PO3AD (Ye et al., 2025)	<u>0.5226</u>	<u>0.5519</u>	<u>0.5700</u>	<u>0.0233</u>	<u>0.6637</u>	<u>0.7370</u>	<u>0.6814</u>	<u>0.1300</u>
R3D-AD (Zhou et al., 2024)†	0.5165	0.5404	0.4852	0.0127	0.5053	0.6009	0.5051	0.0154
NOVA	**0.9127**	**0.8933**	**0.9617**	**0.5213**	**0.9656**	**0.9810**	**0.9729**	**0.7701**

†R3D-AD evaluates 4,096/2,048-point subsets on Real3D-AD/Anomaly-ShapeNet; the other detectors use 16,384 points. Training and sampling details are given in Appendix D.3.

5 Conclusion

We presented NOVA, a variation-driven framework for 3D anomaly synthesis and detection. NOVA derives local deformation patterns from differences among aligned normal surfaces and controls their strength relative to observed normal variation, avoiding predefined anomaly shapes while producing diverse pseudo-anomalies. For detection, NOVA uses geometric features learned from normal samples to score local anomalies. Keeping the normal-geometry model fixed, it trains the anomaly-scoring branch through pointwise comparisons before and after synthetic deformation. Notably, the normal-geometry branch alone already surpasses the best published object- and point-level AUROC results on both benchmarks. With 32 pseudo-anomalies per category, P-AUROC rises from 92.68% to 95.99% on Real3D-AD and from 96.94% to 97.16% on Anomaly-ShapeNet, with O-AUROC reaching 91.28% and 96.37%, respectively. Increasing the synthesis quantity does not consistently yield further gains.

REPRODUCIBILITY STATEMENT

Section 3 describes anomaly synthesis and detection, and Section 4.1 specifies the datasets and evaluation metrics. Appendices A and B provide synthesis parameters, network architectures, training objectives, calibration, and training settings. Appendix C documents baseline and comparison protocols, and Appendix D reports additional experiments, including multi-seed and per-category results. Code and workflow documentation are provided at `https://anonymous.4open.science/r/NOVA-7C92/`.

AI USE STATEMENT

We used generative AI tools to assist with grammar checking, language polishing, and the organization of parts of the code. All AI-assisted text, code, technical claims, and numerical results were reviewed and verified by the authors. The authors take full responsibility for the final content of the paper.

REFERENCES

Pranav A, Shashank B, Pranav Siddappa, Dominik Seuss, Minal Moharir, and Subramanya KN. Two steps are all you need: Efficient 3D point cloud anomaly detection with consistency models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 3479–3487, 2026.

Ali Balapour and Faraz Hach. Anomaly Factory 3D: A modular framework for diverse pseudo-anomaly synthesis in unsupervised 3D anomaly detection. In *Proceedings of the European Conference on Computer Vision*, 2026.

Paul Bergmann and David Sattlegger. Anomaly detection in 3D point clouds using deep geometric descriptors. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2613–2623, 2023.

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.

Yunkang Cao, Xiaohao Xu, and Weiming Shen. Complementary pseudo multimodal feature for point cloud anomaly detection. *Pattern Recognition*, 156:110761, 2024.

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 *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 837–845, 2025a.

Yuqi Cheng, Yunkang Cao, Dongfang Wang, Weiming Shen, and Wenlong Li. Boosting global-local feature matching via anomaly synthesis for multi-class point cloud anomaly detection. *IEEE Transactions on Automation Science and Engineering*, 22:12560–12571, 2025b.

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. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 3327–3334, 2026.

Dinh-Cuong Hoang, Phan Xuan Tan, Anh-Nhat Nguyen, Ta Huu Anh Duong, Tuan-Minh Huynh, Minh-Anh Nguyen, Duc-Manh Nguyen, Minh-Duc Cao, Duc-Huy Ngo, Minh-Quang Vu, Van-Duc Vu, Van-Thiep Nguyen, Thu-Uyen Nguyen, Khanh-Toan Phan, and Van-Hiep Duong. Vote3D-AD: Unsupervised point cloud anomaly localization via varied defect synthesis and differentiable vote-clustering. *Alexandria Engineering Journal*, 137:171–193, 2026a.

Dinh-Cuong Hoang, Phan Xuan Tan, Anh-Nhat Nguyen, Minhhuy Le, Ta Huu Anh Duong, Tuan-Minh Huynh, Duc-Manh Nguyen, Minh-Duc Cao, Duc-Huy Ngo, Minh-Quang Vu, Thu-Uyen Nguyen, Khanh-Toan Phan, Minh-Quang Do, Xuan-Tung Dinh, Van-Hiep Duong, and Van-Thiep Nguyen. Efficient industrial point cloud anomaly detection via spatial context aggregation and selective anomalous feature generation. *Scientific Reports*, 16(1):10309, 2026b.

Dinh-Cuong Hoang, Phan Xuan Tan, Anh-Nhat Nguyen, Duc-Huy Ngo, Minh-Duc Cao, Minh-Quang Vu, Hoang-Nam Duong, Sy-Huong Nguyen, Thi-Hong Le, Van-Viet Dang, Xuan-Tung Dinh, Minh-Anh Nguyen, Minh-Quang Do, Van-Khanh Giap, and Van-Hiep Duong. Multi-scale token fusion with Mamba adapters for efficient 3D industrial defect detection. *Array*, 30:100976, 2026c.

Eliahu Horwitz and Yedid Hoshen. Back to the feature: Classical 3D features are (almost) all you need for 3D anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2968–2977, 2023.

Xueyang Kang, Zizhao Li, Tian Lan, Dong Gong, Kourosh Khoshelham, and Liangliang Nan. Hierarchical point-patch fusion with adaptive patch codebook for 3D shape anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 24258–24267, 2026.

SuYeon Kim, Wongyu Lee, and MyeongAh Cho. A semantically disentangled unified model for multi-category 3D anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 33036–33045, 2026.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Chun-Liang Li, Kihyuk Sohn, Jinsung Yoon, and Tomas Pfister. CutPaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9664–9674, 2021.

Wenqiao Li, Xiaohao Xu, Yao Gu, Bozhong Zheng, Shenghua Gao, and Yingna Wu. Towards scalable 3D anomaly detection and localization: A benchmark via 3D anomaly synthesis and a self-supervised learning network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22207–22216, 2024.

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. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 5146–5154, 2025a.

Hanzhe Liang, Jie Zhang, Tao Dai, Linlin Shen, Jinbao Wang, and Can Gao. Taming anomalies with down-up sampling networks: Group center preserving reconstruction for 3D anomaly detection. In *Proceedings of the ACM International Conference on Multimedia*, pp. 7133–7141, 2025b.

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. In *Advances in Neural Information Processing Systems, Datasets and Benchmarks Track*, volume 36, 2023a.

Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. SimpleNet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20402–20411, 2023b.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Xiaoling Lv, Min Cao, Qiong Xie, and Silong Peng. From comparison to discrimination: A one-step framework for 3D anomaly detection and localization. *Image and Vision Computing*, 172:106032, 2026.

Jian Ning, Qin Zou, Linchun Wu, Yuanhao Yue, Kunmo Li, Shoubin Chen, and Zhongyuan Wang. Physics-inspired pseudo anomaly generation and prototype feature guidance for 3D anomaly detection. *Pattern Recognition*, 180:114391, 2026.

Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas Guibas. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14318–14328, 2022.

Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (FPFH) for 3D registration. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3212–3217, 2009.

Jiaxiang Wang, Haote Xu, Xiaolu Chen, Haodi Xu, Yue Huang, Xinghao Ding, and Xiaotong Tu. Exploiting point-language models with dual-prompts for 3D anomaly detection. *Expert Systems with Applications*, 298:129758, 2026.

Linchun Wu, Jian Ning, and Qin Zou. Anomaly-aware Siamese comparative transformer for 3D anomaly detection. *Pattern Recognition Letters*, 207:137–144, 2026a.

Linchun Wu, Qin Zou, Yuanhao Yue, and Zhongyuan Wang. Geometry-aligned and anomaly-aware reconstruction for 3D anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14648–14657, 2026b.

Heng Yang, Jingnan Shi, and Luca Carlone. TEASER: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2021.

Jianan Ye, Weiguang Zhao, Xi Yang, Guangliang Cheng, and Kaizhu Huang. PO3AD: Predicting point offsets toward better 3D point cloud anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1353–1362, 2025.

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 *Advances in Neural Information Processing Systems*, volume 38, 2025.

Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. DRAEM – a discriminatively trained reconstruction embedding for surface anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8330–8339, 2021.

Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. Keep DRÆMing: Discriminative 3D anomaly detection through anomaly simulation. *Pattern Recognition Letters*, 181:113–119, 2024.

Yaohua Zha, Xue Yuerong, Chunlin Fan, Yuansong Wang, Tao Dai, Ke Chen, and Shu-Tao Xia. CASL: Curvature-augmented self-supervised learning for 3D anomaly detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12340–12348, 2026.

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 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 27063–27072, 2025.

Zheyuan Zhou, Le Wang, Naiyu Fang, Zili Wang, Lemiao Qiu, and Shuyou Zhang. R3D-AD: Reconstruction via diffusion for 3D anomaly detection. In *Proceedings of the European Conference on Computer Vision*, pp. 91–107, 2024.

Hongze Zhu, Guoyang Xie, Chengbin Hou, Tao Dai, Can Gao, Jinbao Wang, and Linlin Shen. Towards high-resolution 3D anomaly detection via group-level feature contrastive learning. In *Proceedings of the ACM International Conference on Multimedia*, pp. 4680–4689, 2024.

A SYNTHESIS DETAILS

A.1 COMPUTING SURFACE DIFFERENCES

Alignment and normalization. Normal training objects define each category's reference pose and normalization. Objects are centered individually and rigidly aligned using FPFH matching, TEASER++, and ICP refinement with multiple initializations. The training reference minimizes mean pairwise registration error. Real3D-AD uses bounding-box midpoints for instance centering; Anomaly-ShapeNet uses the midpoint of the coordinate-wise 0.5th and 99.5th percentiles. After alignment, the category center is the midpoint of pooled training bounds, and the scale is half their largest axis span with padding. The bounds use quantiles 0.005/0.995 with 5% padding on Real3D-AD and 0.001/0.999 with 8% padding on Anomaly-ShapeNet. Test objects are registered to the training references and use the same saved category normalization. Real3D-AD coordinates are clipped to $[-1 + 10^{-6}, 1 - 10^{-6}]$; Anomaly-ShapeNet coordinates remain unclipped. Transform records preserve point-label correspondence. Deterministic farthest-point sampling (FPS) then selects 16,384 points, with the same saved indices used throughout evaluation.

Normals and local differences. Normals are the smallest-eigenvalue directions of the second-moment matrices formed by 24 non-center neighbors relative to each query point. Their signs are propagated over the neighbor graph, aligned across objects to the first normal reference, and corrected by a radial check for an entirely inward orientation. Local spacing ℓ_i is the median distance to these 24 neighbors.

A point in the normal object X^s is matched to its nearest point $j_t(i)$ in target X^t. A match is valid only if it is mutual and the oriented normal dot product is at least 0.5; let $\gamma_i^{s\to t}$ indicate validity. The signed offset along the source normal to the target tangent plane, in local-spacing units, is

$$r_i^{s\to t} = \begin{cases} \dfrac{(x^t_{j_t(i)} - x_i^s)^\top n^t_{j_t(i)}}{\max(\ell_i^s, 10^{-6})\,(n_i^s)^\top n^t_{j_t(i)}}, & \gamma_i^{s\to t} = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

The local region Ω_c contains the 512 closest reachable points to center c on a symmetrized 16-neighbor graph with Euclidean edge lengths; smaller connected components are rejected. With graph distance d_i and $R_c = \max(\max_{i\in\Omega_c} d_i, 10^{-12})$, the boundary weight is $w_i = [1-(d_i/R_c)^2]_+^2$, where $[\cdot]_+$ is the positive part. To separate broad surface offsets from local variation, we fit a weighted ridge trend with coefficients $\theta \in \mathbb{R}^4$ in tangent-plane coordinates (ξ_i, η_i):

$$\theta^{s\to t} = \arg\min_\theta \sum_{i\in\Omega_c} \gamma_i^{s\to t} w_i \big(r_i^{s\to t} - \psi_i^\top\theta\big)^2 + \lambda\|\theta\|_2^2, \quad \psi_i = [1, \xi_i, \eta_i, \xi_i^2 + \eta_i^2]^\top. \tag{A.2}$$

The residual used for shape extraction is $h_i^{s\to t} = \gamma_i^{s\to t}(r_i^{s\to t} - \psi_i^\top\theta^{s\to t})$.

A.2 DEFORMATION CONSTRUCTION AND PARAMETERS

For each of the T aligned normal targets, residual normalization is

$$\bar{h}_i^{(t)} = \frac{h_i^{s\to t}}{\max(Q_{0.95}(\{|h_i^{s\to t}|\}; \{\gamma_i^{s\to t} w_i\}), \varepsilon)}. \tag{A.3}$$

Here $Q_\tau(\{z_i\}; \{\omega_i\})$ is the weighted empirical quantile: the first sorted value whose cumulative weight reaches τ times the total weight (the total is floored at ε); omitted weights are equal.

The combined pattern p_i is given by Equation 1. Percentile ranks $\rho_i^{(t)}$ assign average ranks to ties and are rescaled to $[0, 1]$. Let v_i denote the numerator of Equation 1, set to zero at points without common valid matches or positive weight. The normalizer is $Z = \max(Q_{0.95}(\{|v_i|\}; \{w_i\}), \varepsilon)$.

The normal amplitude reference is computed from surface offsets before trend removal:

$$b_{\mathcal{C}} = Q_{0.99}\Big(\big\{Q_{0.95}(\{|r_i^{s\to t}|\}; \{\gamma_i^{s\to t} w_i\})\big\}_{c,s,t}\Big). \tag{A.4}$$

The inner quantile is over neighborhood points; the outer quantile covers source–target–center combinations. Table 5 lists the settings.

Table 5: **Anomaly synthesis settings.** Both datasets share the construction settings and use 32 training pairs per category in the main experiment.

Setting	Value
Sampled points; normal/spacing neighbors	16,384; 24 (excluding center)
Surface-graph neighbors; region size	16; 512 points
Normal matching; valid support weight	Normal dot product ≥ 0.5; $\geq 20\%$ per target
Trend regularization; numerical floor	$\lambda = 10^{-3}$; $\varepsilon = 10^{-8}$
Residual-amplitude and normal-reference quantiles	Weighted 95th percentile; event-wise 99th percentile
Added amplitudes Δ_e	$0.5,\ 1.0,\ 1.5$ (subtle, medium, strong)
Normal objects; targets per source	4; the other 3 normal objects
Main training pairs (Real3D-AD / Anomaly-ShapeNet)	32 / 32
Matched-control training pairs	16 / 32

Each training pair contains one normal object and its deformed counterpart at a sampled center and strength. Centers are considered in fixed-seed random order and require at least 20% valid support weight per target.

A.3 Synthesis Examples Across Categories

Figure 6 applies the same synthesis procedure to additional objects from Real3D-AD and Anomaly-ShapeNet.

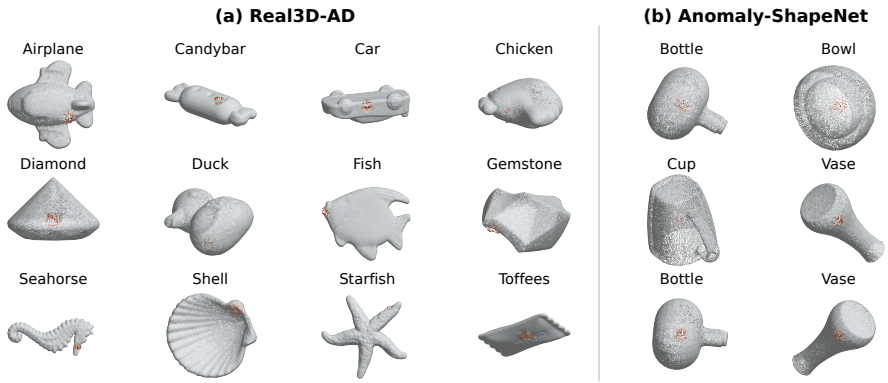


Figure 6: **Anomaly synthesis across object categories.** The same procedure generates local anomalies on (a) Real3D-AD and (b) Anomaly-ShapeNet. Warmer colors indicate larger point displacement.

B Detector Details

The encoder and covariance predictor learn from normal data; synthetic pairs update only the anomaly-scoring head.

B.1 Local Encoding and Covariance Prediction

Neighborhood representation. The encoder uses 15 neighbors excluding the center point. Let μ_i and M_i be their centroid and centered second-moment matrix, and define $S_i^{\mathrm{loc}} = M_i / \max(\mathrm{tr}\, M_i, \varepsilon)$. A shared radial MLP embeds $\log(1 + \|x_j - \mu_i\|_2^2)$ for each neighbor. Mean, standard-deviation, and maximum pooling give 48 channels; adding $\log(1 + \mathrm{tr}\, M_i)$ and $\|S_i^{\mathrm{loc}}\|_F$ forms the 50-dimensional input to the 32-dimensional descriptor encoder. Table 6 gives the layer dimensions.

Conditional covariance network. The query and matched-reference descriptors contribute 64 input channels. Eight additional scalars summarize the two log-traces and shape Frobenius norms, the query–reference shape inner product, the log-squared centroid gap, and the two shape inner products with the centroid-gap matrix. In Equation B.1, S_i^q and S_{ir}^r are the corresponding trace-normalized matrices S^{loc}. With centroid difference $g_{ir} = \mu_i^q - \mu_{ir}^r$, the remaining matrix is $S_{ir}^g = g_{ir}g_{ir}^\top/(\|g_{ir}\|_2^2 + \varepsilon)$. The resulting 72-dimensional input is mapped to a 64-dimensional matching feature and then four outputs $(\beta_0, \beta_q, \beta_r, \beta_g)$. The covariance is

$$\Sigma_{ir} = \lambda_{\mathrm{iso}} I + \lambda_q S_i^q + \lambda_r S_{ir}^r + \lambda_g S_{ir}^g, \tag{B.1}$$

where I is the identity matrix and the weights are positive. The shape matrices describe the directional spread of the two neighborhoods and their centroid difference; their weights are parameterized as

$$\begin{aligned} \lambda_{\mathrm{iso}} &= \sigma_0^2 \exp(2 \log 4 \, \tanh \beta_0), \\ \lambda_a &= 8\sigma_0^2 \, \mathrm{sigmoid}(\beta_a), \qquad a \in \{q, r, g\}. \end{aligned} \tag{B.2}$$

We use $\varepsilon = 10^{-8}$ and a fixed scale $\sigma_0 = 0.0062600793$ in normalized-coordinate units. This scale is estimated from the minimum nearest-neighbor distance to the four normal references for every sixteenth point of the Real3D-AD clean and scan-perturbed training samples. Distances are pooled across categories, and their 75th percentile is taken with a floor of 10^{-4}. Anomaly-ShapeNet uses the same scale with independently initialized learnable weights. The positive isotropic term and positive-semidefinite shape terms keep the covariance positive definite.

Normal fitting. Let $j_r(i)$ select the nearest reference point, $d_{ir} = x_{j_r(i)}^r - x_i$, and $\mathcal{R}_i$ contain the valid references. The Gaussian energy and mixture negative log likelihood are

$$\begin{aligned} e_{ir} &= \tfrac{1}{2}(d_{ir}^\top \Sigma_{ir}^{-1} d_{ir} + \log|\Sigma_{ir}| + 3\log(2\pi)), \\ \mathcal{L}_{\mathrm{NG}} &= -\frac{1}{N}\sum_i \log\left[\frac{1}{M_{\mathcal{C}}}\sum_{r\in\mathcal{R}_i} \exp(-e_{ir})\right]. \end{aligned} \tag{B.3}$$

Here N is the query point count and $M_{\mathcal{C}} = 4$ the reference count. The mixture weight remains $1/M_{\mathcal{C}}$ when invalid references are masked. All four references are used, except that Anomaly-ShapeNet excludes a clean query's own reference during normal fitting and calibration.

Table 6: **Network architecture and parameter counts.** The architecture differs only in the scoring-head input width. Paired values denote Real3D-AD/Anomaly-ShapeNet; counts include biases and normalization parameters.

Component	Mapping	Parameters
Local encoder	Radial $1 \to 16 \to 16$; pooled $50 \to 64 \to 32$, final LayerNorm	5,712
Matching features	$72 \to 64 \to 64$	8,832
Covariance gates	$64 \to 4$	260
Anomaly-scoring head	$D \to 64 \to 64 \to 1$; $D = 233/236$	19,201 / 19,393

Arrows denote affine layers, with SiLU after hidden layers except where LayerNorm is specified. The normal-geometry branch has 14,804 trainable parameters; the fixed σ_0 buffer is excluded.

B.2 Anomaly-Scoring Head Inputs

Table 7 lists the head's inputs from the frozen normal-geometry branch, including local descriptors and reference-wise summaries.

Table 7: **Inputs to the anomaly-scoring head.** Features come from the frozen normal-geometry branch; datasets differ in the displacement channels. Mean/max/std summarize references, and zero marks an omitted group.

Feature group	Real3D-AD	Anomaly-ShapeNet
Query local geometry descriptor	32	32
Mean/max/std of 64-D matching features	192	192
Mean/max/std of Mahalanobis statistics	3	3
Mean/max/std of covariance log-determinants	3	3
Mean/max/std of displacement magnitudes	0	3
Reference-agreement statistics	3	3
Total	233	236

Reference agreement is measured from the Gaussian-mixture posterior over references. On Real3D-AD, we omit one reference at a time and summarize the posterior entropy by its mean, maximum, and population standard deviation. The geometry, Mahalanobis, and log-determinant summaries are computed within each three-reference subset and averaged across the four subsets; direct displacement magnitudes are excluded. Anomaly-ShapeNet uses the full valid bank, with normalized posterior entropy, maximum posterior probability, and posterior population standard deviation as its three agreement channels, plus three summaries of $\|d_{ir}\|_2$.

Inputs are standardized using clean and scan-perturbed normal samples. The MLP has a zero-initialized output layer, no dropout, and no additional normalization; it is trained with Equation 3.

B.3 Score Calibration and Spatial Aggregation

In Equation 4, raw-score quantiles are estimated per category from clean and scan-perturbed normal samples after head training. These fixed quantiles calibrate all test objects of that category.

The spatial operator $\mathcal{A}$ in Equation 5 uses two rounds of neighbor averaging:

$$\mathcal{G}_{16}(v)_i = \mathrm{Mean}_{j\in\mathcal{N}_{16}(i)}\,\mathrm{Mean}_{k\in\mathcal{N}_{16}(j)}\,v_k. \tag{B.4}$$

$\mathcal{N}_{16}(i)$ includes the query point, and $\mathcal{A}(v)_i = \mathrm{softplus}(\mathcal{G}_{16}(v)_i)$. NOVA uses $\mathcal{A}(E)$ for point localization and its maximum for object scoring.

Normal-only comparison. The Normal-only output uses the same frozen normal-geometry model without the anomaly-scoring head. For valid references $\mathcal{R}_i$, it takes the minimum standardized geometric deviation,

$$D_i^M = \min_{r\in\mathcal{R}_i} d_{ir}^\top \Sigma_{ir}^{-1} d_{ir}, \tag{B.5}$$

and transforms it to $T_i^N = -\log \overline{F}_{\chi_3^2}(D_i^M)$, where $\overline{F}_{\chi_3^2}$ is the chi-square survival function with three degrees of freedom. Let q_p^N be the pth percentile of T^N on normal calibration data. The calibrated geometric response is

$$z_i^N = \frac{T_i^N - q_{99}^N}{\max(q_{99}^N - q_{90}^N, 10^{-6})}. \tag{B.6}$$

Geometric quantiles are pooled across training categories and fixed before head training. Real3D-AD uses held-out normal variants after initial likelihood fitting; Anomaly-ShapeNet uses clean and scan-perturbed normal samples after normal fitting. The chi-square transform places geometric deviations on a common reference scale. Normal-only uses $\mathcal{A}(z^N)$ for localization and its maximum for object scoring.

B.4 Training Settings

The head is trained for eight epochs with the normal-geometry branch frozen. Both stages use AdamW (Loshchilov & Hutter, 2019) with initial learning rate 3×10^{-4}. Within each dataset, the normal-geometry model and scoring head each share parameters across categories. Synthesis strengths are listed in Table 5; seed variability is reported in Appendix D.1.

Real3D-AD uses 768 likelihood-fitting updates followed by the normal-only refinement below; Anomaly-ShapeNet uses likelihood fitting alone for eight epochs of 192 updates. Head-training seeds vary the head initialization and sample order, with the normal-geometry model and synthesis pool fixed. For NOVA, shaded bands and $\pm$ values indicate sample standard deviations across head-training seeds.

B.5 Real3D-AD Normal-Only Refinement

Before freezing the Real3D-AD normal-geometry branch, we refine it to suppress high responses on normal surfaces while retaining their likelihood. This stage uses only clean and scan-perturbed normal samples, with all four references. Training mixes groups of two clean queries and one scan-perturbed query with scan-only groups. Each group contributes equally, and gradients from eight groups are accumulated per optimizer update.

Let g_i denote the learned Gaussian-mixture log likelihood minus that of the same reference mixture with fixed isotropic covariance $\sigma_0^2 I$. Let $\hat{z}_i$ be the calibrated geometric response in Equation B.6, computed with the current normal model and quantiles fixed after the initial likelihood-fitting stage. For a normal query X, let $\hat{o}(X)$ be the mean of its largest 1% of point responses $\mathrm{softplus}(\hat{z}_i)$, before neighborhood aggregation. With $h(v) = 0.1\,\mathrm{softplus}(v/0.1)$, we minimize

$$\mathcal{L}_{\mathrm{refine}} = \mathrm{Mean}_i[0.05 - g_i]_+ + 0.25\,\mathrm{Mean}_i\, h(\hat{z}_i) + \mathrm{Mean}_X\, h(\hat{o}(X) - 1). \tag{B.7}$$

Here $[v]_+ = \max(v, 0)$. The pointwise means include all points of the normal queries in each group, and the last mean is over those queries. The first term preserves likelihood gain; the remaining terms penalize high point responses and their upper tail within each object.

Refinement updates both the local encoder and covariance predictor for 1,536 steps (eight epochs of 192 updates), using a fresh AdamW optimizer with learning rate $3{\times}10^{-4}$, weight decay 10^{-4}, cosine learning-rate decay, and gradient clipping at 1. The normal-model seed is fixed at 20260812. The calibration quantiles remain fixed throughout refinement and subsequent head training. At 16,384 points, the top-1% training statistic averages 164 responses. At inference, the object score is the maximum of the spatially aggregated point map (Appendix B.3).

C Baseline and Comparison Protocols

C.1 Published Results in Table 1

Table 1 reproduces baseline results reported in the cited papers, their supplementary materials, and published benchmark comparisons (Liu et al., 2023a; Li et al., 2024; Kim et al., 2026). A dash indicates that the selected source provides no result for the category set evaluated here. CPMF's Real3D-AD results follow its published journal version. DUS-Net uses the main results, which differ from its abstract. AARD's Anomaly-ShapeNet O-AUROC is retained as reported, although its correspondence to the standard 40-category average is unspecified.

Over three seeds, PO3AD with AF3AD synthesis reports standard deviations of 0.012, 0.019, 0.021, and 0.016 for the four columns of Table 1. NOVA's corresponding values are given in Table 8 at 32 pairs.

C.2 Implementation of Comparison Synthesis Methods

Table 2 uses four synthesis strategies adapted to the same aligned 16,384-point normal objects. Normals are estimated from 24 neighbors. Each generator operates in its native centered, rotated, and normalized coordinates; its displacement is then mapped back to the original coordinates, preserving point indices and leaving unselected points unchanged. Each source provides 16/32 pseudo-anomalies per Real3D-AD/Anomaly-ShapeNet category, with the NOVA detector and head-training settings fixed.

On Anomaly-ShapeNet, normal-model fitting, head training, and calibration use the same 40 categories as evaluation. All sources share the frozen normal-geometry model, 1,312 test objects, point sampling, labels, and three head-training seeds. Metrics are macro-averaged over categories within each seed.

PO3AD. Following the normal-guided displacement formulation of Ye et al. (2025), we partition the object into 64 patches and select one from the first half of the patch list, as in the released sampler. Points in that patch move along their normals or the opposite direction, with a common random sign. Displacement decreases linearly with distance from the patch center and is rescaled to $[0, 1]$ before multiplication by a magnitude sampled uniformly from $[0.06, 0.12]$. Our implementation of the published equations uses signed normals in place of the coordinatewise absolute normals used by the released implementation.

CASL. CASL adopts the same normal-guided local displacement strategy (Zha et al., 2026), with magnitude sampled uniformly from $[0.06, 0.12]$ and signed normals in our implementation of the published equations. The main difference is region selection. On Real3D-AD, the released sampler first selects one side of a random plane and draws a patch count $J \in \{10, \ldots, 100\}$ with probability proportional to $\exp[-(1/J - 0.05)^2/(2 \cdot 0.015^2)]$. On Anomaly-ShapeNet it uses 100 patches. One patch is selected from the full patch list; its displacement is restored to the original complete point cloud.

AF3AD. We use the released modular synthesizer of Balapour & Hach (2026) with the PO3AD integration's 64-patch partition. A patch is selected from the full list, and one of 11 deformation presets is sampled uniformly. These include bulges, dents, ridges, trenches, shear, ripples, and directional stretching. Radius and magnitude use uniform ranges $[0.03, 0.25]$ and $[0.06, 0.12]$, respectively; the micro-dimple preset applies five successive local deformations. Outputs whose final coordinates are exactly unchanged are redrawn, so this comparison samples AF3AD conditional on nonzero deformation. The retained pools contain 192 accepted samples from 213 attempts on Real3D-AD and 1,280 from 1,358 attempts on the original 40 Anomaly-ShapeNet categories.

PA3AD. We independently implement the paper equations of Ning et al. (2026): normal-directed Gaussian bulges and concavities, signed displacements along the leading local principal axis for cracks, and weighted bending displacements around an axis. Hole generation is excluded to preserve point correspondence. The four retained types have probabilities $7/18$, $7/18$, $1/9$, and $1/9$. Magnitude is uniform on $[0.08, 0.24]$. We also sample Gaussian width independently from this interval, since its distribution is unspecified. The released region selector uses size 0.05. Bending follows the leading local principal axis with the released radial window, constant inside radius r and quadratically tapered to zero at $1.5r$; its denominator uses a 10^{-8} stabilizer.

C.3 Qualitative Comparison Protocol

Samples and correspondence. Released CASL/PA3AD checkpoints retain their native preprocessing and scorers. NOVA uses 32 pairs per category and the first head-training seed, with 40 Anomaly-ShapeNet categories for training and evaluation. Scores are mapped to the same 16,384 points for metric computation and rendering, using identical views and depth ordering. The 12 selected examples comprise two anomalies each from Shell and Duck for CASL and one anomalous/normal pair each from Ashtray, Bag, Bottle, and Vase for PA3AD. They illustrate localization improvements and, for PA3AD, reduced false positives. Category-average performance is reported separately; all evaluated candidates are retained in the experiment records.

Display and calibration. CASL colors show within-object relative intensity, using linear min–max scaling for each method; point AP uses raw scores. PA3AD panels use fixed category-wise scales: four training-normal objects determine each method's median q_{50} and 99th percentile q_{99}. NOVA excludes a calibration query from its own reference bank and uses all references for test objects. Quantiles use the higher order statistic. Colors show $\mathrm{clip}((s - q_{50})/\max(q_{99} - q_{50}, 10^{-8}), 0, 3)$ for point score s, and point FPR is the fraction of ground-truth normal points with $s > q_{99}$. All panels display every point.

D Additional Results

D.1 Complete Results of the Synthesis Quantity Study

Table 8 gives the complete metrics for Figure 5, using nested pools of 8, 16, 32, 64, and 96 pseudo-anomalies per category and the same 40 Anomaly-ShapeNet categories for normal-geometry training and evaluation.

Table 8: **Effect of synthesis quantity on detection.** Under eight-epoch head training, 96 pairs yield lower mean O-AUROC than 64 on both datasets. Shading marks the 32-pair main setting on both datasets. The best and second-best means are shown in bold and underlined, respectively.

Dataset	Count	O-AUROC	O-AP	P-AUROC	P-AP
Real3D-AD	8	$\underline{0.9153} \pm 0.0030$	0.8937 ± 0.0024	$\underline{0.9599} \pm 0.0009$	0.5181 ± 0.0020
	16	$\mathbf{0.9155} \pm 0.0035$	$\underline{0.8957} \pm 0.0023$	$\mathbf{0.9616} \pm 0.0007$	$\mathbf{0.5249} \pm 0.0055$
	32	0.9128 ± 0.0012	$\mathbf{0.8959} \pm 0.0005$	$\underline{0.9599} \pm 0.0006$	$\underline{0.5218} \pm 0.0074$
	64	0.9046 ± 0.0009	0.8907 ± 0.0015	0.9547 ± 0.0013	0.4916 ± 0.0048
	96	0.9008 ± 0.0020	0.8862 ± 0.0013	0.9537 ± 0.0017	0.4889 ± 0.0093
Anomaly-ShapeNet	8	0.9633 ± 0.0010	0.9798 ± 0.0006	0.9724 ± 0.0002	$\underline{0.7689} \pm 0.0019$
	16	0.9631 ± 0.0014	0.9797 ± 0.0007	0.9723 ± 0.0009	0.7684 ± 0.0008
	32	$\underline{0.9637} \pm 0.0018$	0.9800 ± 0.0009	0.9716 ± 0.0012	0.7682 ± 0.0017
	64	$\mathbf{0.9644} \pm 0.0011$	$\mathbf{0.9806} \pm 0.0004$	$\mathbf{0.9733} \pm 0.0003$	$\mathbf{0.7704} \pm 0.0007$
	96	0.9631 ± 0.0008	$\underline{0.9801} \pm 0.0005$	$\underline{0.9730} \pm 0.0010$	0.7676 ± 0.0028

D.1.1 Training Budget and Scoring-Head Width

The eight-epoch quantity study changes both the synthesis pool size and the number of optimizer updates. We examine these effects using 32, 64, and 96 pairs per category, two scoring-head widths, and training durations up to 32 epochs. Tables 9 and 10 report all evaluated configurations on Real3D-AD and Anomaly-ShapeNet.

Controlled settings. The normal-geometry model, nested synthesis pools, features, loss, optimizer, and calibration remain fixed. We compare two-hidden-layer MLPs of width 64 and 128. Each configuration is trained from scratch for 32 epochs with three head-training seeds and evaluated at epochs 8, 16, and 32. Additional checkpoints at epoch 24 for 32 pairs and epoch 12 for 64 pairs give equal-update comparisons: $32 \times 24 = 64 \times 12 = 96 \times 8 = 768$ updates per category. Each update processes one normal/anomaly/scan-perturbed group. The tables include all 132 evaluations from 36 training runs, using Equation 5 and the main dataset splits. Metrics are averaged over categories within each seed, then over the three head-training seeds.

Equal-update comparison. With width 64 and matched updates, Real3D-AD O-AUROC is 0.8993, 0.8997, and 0.9008 for 32, 64, and 96 pairs, compared with 0.9128, 0.9046, and 0.9008 after eight epochs. Width 128 gives similarly close equal-update values of 0.8997, 0.8993, and 0.8994. Anomaly-ShapeNet shows metric-dependent differences at equal updates. Thus, the quantity trend depends on the training budget.

Longer training and a wider head. Longer training and a wider head give mixed results across metrics and datasets. For 32 pairs and width 64, extending training from eight to 32 epochs lowers Real3D-AD O-AUROC from 0.9128 to 0.8978. The same change raises Anomaly-ShapeNet P-AUROC from 0.9716 to 0.9732 while lowering O-AUROC from 0.9637 to 0.9624. Width 128 also provides no consistent improvement. The main comparisons retain 32 pairs, width 64, and eight epochs.

Table 9: **Synthesis quantity, training duration, and scoring-head width on Real3D-AD.** Shading marks the main setting; † marks equal-update configurations (768 updates per category). Matching updates reduces O-AUROC differences across sample counts; longer training and wider heads provide no consistent improvement.

Pairs	Width	Epochs	O-AUROC	O-AP	P-AUROC	P-AP
32	64	8	0.9128 ± 0.0012	0.8959 ± 0.0005	0.9599 ± 0.0006	0.5218 ± 0.0074
32	64	16	0.9037 ± 0.0005	0.8891 ± 0.0005	0.9565 ± 0.0016	0.4996 ± 0.0092
32†	64	24	0.8993 ± 0.0017	0.8842 ± 0.0008	0.9485 ± 0.0053	0.4867 ± 0.0060
32	64	32	0.8978 ± 0.0006	0.8836 ± 0.0004	0.9509 ± 0.0006	0.4802 ± 0.0015
64	64	8	0.9046 ± 0.0009	0.8907 ± 0.0015	0.9547 ± 0.0013	0.4916 ± 0.0048
64†	64	12	0.8997 ± 0.0037	0.8852 ± 0.0003	0.9495 ± 0.0034	0.4850 ± 0.0137
64	64	16	0.8976 ± 0.0014	0.8836 ± 0.0021	0.9497 ± 0.0055	0.4777 ± 0.0059
64	64	32	0.8974 ± 0.0026	0.8810 ± 0.0028	0.9498 ± 0.0036	0.4808 ± 0.0016
96†	64	8	0.9008 ± 0.0020	0.8862 ± 0.0013	0.9537 ± 0.0017	0.4889 ± 0.0093
96	64	16	0.8961 ± 0.0029	0.8803 ± 0.0033	0.9490 ± 0.0048	0.4746 ± 0.0162
96	64	32	0.8983 ± 0.0039	0.8801 ± 0.0026	0.9544 ± 0.0029	0.4907 ± 0.0160
32	128	8	0.9088 ± 0.0019	0.8933 ± 0.0017	0.9583 ± 0.0012	0.5124 ± 0.0046
32	128	16	0.9038 ± 0.0016	0.8893 ± 0.0018	0.9560 ± 0.0018	0.5004 ± 0.0073
32†	128	24	0.8997 ± 0.0036	0.8832 ± 0.0013	0.9491 ± 0.0052	0.4937 ± 0.0036
32	128	32	0.8999 ± 0.0022	0.8853 ± 0.0020	0.9506 ± 0.0022	0.4881 ± 0.0065
64	128	8	0.9030 ± 0.0015	0.8890 ± 0.0007	0.9534 ± 0.0029	0.4849 ± 0.0085
64†	128	12	0.8993 ± 0.0056	0.8835 ± 0.0026	0.9469 ± 0.0054	0.4834 ± 0.0166
64	128	16	0.8985 ± 0.0013	0.8825 ± 0.0011	0.9515 ± 0.0065	0.4835 ± 0.0064
64	128	32	0.8992 ± 0.0039	0.8811 ± 0.0048	0.9522 ± 0.0052	0.4949 ± 0.0009
96†	128	8	0.8994 ± 0.0025	0.8850 ± 0.0021	0.9537 ± 0.0021	0.4845 ± 0.0141
96	128	16	0.8970 ± 0.0018	0.8799 ± 0.0013	0.9502 ± 0.0028	0.4782 ± 0.0108
96	128	32	0.8958 ± 0.0025	0.8786 ± 0.0014	0.9546 ± 0.0017	0.4919 ± 0.0073

Table 10: **Synthesis quantity, training duration, and scoring-head width on Anomaly-ShapeNet.** Shading marks the main setting; † marks equal-update configurations (768 updates per category). Longer training can improve point AUROC while reducing object AUROC; a wider head gives mixed results.

Pairs	Width	Epochs	O-AUROC	O-AP	P-AUROC	P-AP
32	64	8	0.9637 ± 0.0018	0.9800 ± 0.0009	0.9716 ± 0.0012	0.7682 ± 0.0017
32	64	16	0.9632 ± 0.0011	0.9800 ± 0.0006	0.9727 ± 0.0001	0.7726 ± 0.0022
32†	64	24	0.9619 ± 0.0012	0.9795 ± 0.0007	0.9728 ± 0.0008	0.7678 ± 0.0030
32	64	32	0.9624 ± 0.0015	0.9796 ± 0.0007	0.9732 ± 0.0005	0.7705 ± 0.0045
64	64	8	0.9644 ± 0.0011	0.9806 ± 0.0004	0.9733 ± 0.0003	0.7704 ± 0.0007
64†	64	12	0.9625 ± 0.0009	0.9797 ± 0.0005	0.9727 ± 0.0008	0.7702 ± 0.0016
64	64	16	0.9637 ± 0.0010	0.9803 ± 0.0003	0.9737 ± 0.0005	0.7697 ± 0.0048
64	64	32	0.9629 ± 0.0012	0.9800 ± 0.0004	0.9733 ± 0.0005	0.7683 ± 0.0055
96†	64	8	0.9631 ± 0.0008	0.9801 ± 0.0005	0.9730 ± 0.0010	0.7676 ± 0.0028
96	64	16	0.9627 ± 0.0012	0.9796 ± 0.0008	0.9730 ± 0.0008	0.7657 ± 0.0081
96	64	32	0.9593 ± 0.0025	0.9783 ± 0.0013	0.9731 ± 0.0007	0.7638 ± 0.0020
32	128	8	0.9635 ± 0.0026	0.9801 ± 0.0013	0.9719 ± 0.0013	0.7695 ± 0.0018
32	128	16	0.9616 ± 0.0025	0.9792 ± 0.0008	0.9726 ± 0.0003	0.7670 ± 0.0105
32†	128	24	0.9622 ± 0.0017	0.9797 ± 0.0012	0.9730 ± 0.0005	0.7674 ± 0.0039
32	128	32	0.9626 ± 0.0025	0.9799 ± 0.0009	0.9733 ± 0.0002	0.7699 ± 0.0024
64	128	8	0.9638 ± 0.0016	0.9803 ± 0.0008	0.9732 ± 0.0006	0.7662 ± 0.0043
64†	128	12	0.9619 ± 0.0003	0.9795 ± 0.0002	0.9723 ± 0.0010	0.7608 ± 0.0041
64	128	16	0.9633 ± 0.0014	0.9803 ± 0.0007	0.9735 ± 0.0001	0.7605 ± 0.0083
64	128	32	0.9620 ± 0.0013	0.9797 ± 0.0005	0.9736 ± 0.0005	0.7678 ± 0.0031
96†	128	8	0.9625 ± 0.0002	0.9800 ± 0.0003	0.9730 ± 0.0010	0.7659 ± 0.0024
96	128	16	0.9609 ± 0.0009	0.9791 ± 0.0006	0.9730 ± 0.0007	0.7604 ± 0.0064
96	128	32	0.9572 ± 0.0013	0.9773 ± 0.0006	0.9732 ± 0.0008	0.7655 ± 0.0084

D.2 Test-Time Sensitivity to Alignment Errors

NOVA compares each query with aligned normal references, so registration errors can change its responses. We estimate residual transforms on normal training data using the registration core with FPFH features (Rusu et al., 2009), TEASER++ global registration (Yang et al., 2021), fallback initializations, and ICP refinement (Besl & McKay, 1992). We replay these transforms on all 1,206 Real3D-AD and 1,312 Anomaly-ShapeNet test objects, covering all 12 and the original 40 categories, respectively. We use the main models trained with 32 synthetic pairs per category and three head-training seeds on both datasets; the Anomaly-ShapeNet model is trained only on those 40 categories. Scoring follows Equation 5: the anomaly map is $\mathcal{A}(E)$ and the object score is its maximum.

Residual estimation. For each category, three normal training objects are moved by a fixed random rigid transform (a 15° rotation and a translation of 2% of the bounding-box diagonal) and registered again. Three types of residual are measured for each object: *self*, registering the transformed object back to its own scan; *cross-instance*, registering the untransformed object to a different normal object; and *repeatability*, the difference between the cross-instance registrations with and without the initial transform. Each residual is applied to every test object of the category, with point sampling, labels, normal references, calibration, and scoring unchanged. All residuals are retained, including poor fits. We average over categories and the three residual replicates within each head-training seed, then summarize across seeds.

Results. Table 11 reports class-macro changes. Median rotation/RMS displacement for self, cross-instance, and repeatability are respectively 0.027°/0.014%, 0.044°/0.027%, and 0.054°/0.035% on Real3D-AD; on Anomaly-ShapeNet they are 3.85°/1.51%, 0.097°/0.058%, and 1.96°/0.76%. RMS is normalized by the object bounding-box diagonal. Small medians do not exclude poor fits: some residual rotations approach 180°.

The largest setting-averaged decreases occur in P-AP, from 0.5218 to 0.5148 on Real3D-AD under cross-instance residuals and from 0.7682 to 0.7601 on Anomaly-ShapeNet under repeatability residuals. Individual categories are more sensitive: the largest absolute P-AP losses for a single residual and head-training seed are 0.1351 for Chicken (cross-instance) and 0.2999 for cap4 (self). Thus, the modest dataset-average changes coexist with substantial localization losses on some categories.

This experiment isolates test-time sensitivity, keeping the synthesis pool and trained models fixed. Self-recovery uses known-pose ground truth; cross-instance transforms also reflect shape differences. The replay measures responses to these specified residuals, whose distribution may differ from actual test-set registration errors. Sensitivity of synthesis to misaligned normal surfaces remains a separate question.

Table 11: **Test-time alignment sensitivity of NOVA.** Fixed main models are evaluated under registration residuals estimated from normal training objects. Unperturbed rows report AUROC/AP; other rows report absolute changes. Dataset-average changes are small, with larger losses on individual categories.

Dataset	Setting	O-AUROC	O-AP	P-AUROC	P-AP
Real3D-AD	Unperturbed	0.9128 ± 0.0012	0.8959 ± 0.0005	0.9599 ± 0.0006	0.5218 ± 0.0074
	Self	$+0.0002 \pm 0.0002$	$+0.0001 \pm 0.0002$	$+0.0002 \pm 0.0001$	$+0.0001 \pm 0.0000$
	Cross-instance	-0.0034 ± 0.0004	-0.0024 ± 0.0002	-0.0031 ± 0.0003	-0.0070 ± 0.0002
	Repeatability	-0.0007 ± 0.0005	-0.0004 ± 0.0003	0.0000 ± 0.0000	$+0.0005 \pm 0.0002$
Anomaly-ShapeNet	Unperturbed	0.9637 ± 0.0018	0.9800 ± 0.0009	0.9716 ± 0.0012	0.7682 ± 0.0017
	Self	-0.0041 ± 0.0009	-0.0022 ± 0.0004	$+0.0006 \pm 0.0002$	-0.0053 ± 0.0004
	Cross-instance	-0.0001 ± 0.0003	0.0000 ± 0.0002	$+0.0004 \pm 0.0002$	-0.0010 ± 0.0001
	Repeatability	-0.0024 ± 0.0010	-0.0013 ± 0.0004	-0.0003 ± 0.0001	-0.0081 ± 0.0004

Each setting uses 36/120 residuals across the 12/40 Real3D-AD/Anomaly-ShapeNet categories. Values are rounded to four decimal places.

D.3 Detector Training with Fixed Synthetic Data

All detectors in Table 4 use the same NOVA-generated anomalies: 16/32 pairs per Real3D-AD/Anomaly-ShapeNet category. This experiment compares how complete detectors use this shared synthetic data, with the training and sampling settings given below. Results use one training seed and each detector's own training and scoring procedures. PointNet++ (Qi et al., 2017) is supervised by normal and anomalous point labels.

PO3AD and R3D-AD. PO3AD uses its offset-prediction objective with 1,000 training epochs (12,000 updates per category), pointwise offset L_1 scores, and the mean point score for object detection. Evaluation covers all 12 Real3D-AD and the original 40 Anomaly-ShapeNet categories, with 1,206 and 1,312 test objects, respectively, at 16,384 points per object. R3D-AD uses batch size 128, Adam with learning rate 10^{-3}, and 40,000 updates, with its relative-reconstruction loss and 200-step diffusion decoder unchanged. It scores distances to reconstructed points and uses its reweighted maximum for object detection on the same test objects. R3D-AD keeps its respective 4,096- and 2,048-point resolutions: deterministic per-object subsets are drawn from the 16,384-point grids, and the corresponding labels are used for point evaluation. For both detectors, AUROC and AP are recomputed from the saved point and object scores with category-wise macro averaging.

PointNet++ training. The baseline is a category-specific PointNet++ with four set-abstraction and four feature-propagation stages, XYZ-only inputs, batch normalization, dropout 0.5, and a binary point-logit output. It is trained from random initialization with Adam (Kingma & Ba, 2015) (learning rate 10^{-3}, weight decay 10^{-4}); each batch contains four groups of anomalous, clean, and scan-perturbed normal clouds at 16,384 points. The training duration is selected on held-out queries from the highest-indexed normal object of each category, by minimizing a class-balanced validation BCE with validation every four epochs (32–200 epochs, plateau learning-rate halving, early stopping after ten unsuccessful checks). Scans are assigned to normal objects by median nearest-neighbor distance, and ambiguous scans are excluded during selection. The final model is retrained on all normal objects, scans, and pseudo-anomalies for the selected duration, which ranges from 32 to 68 epochs on Real3D-AD and 32 to 76 on Anomaly-ShapeNet.

D.4 Per-Category Results

Tables 12 and 13 give per-category Normal-only and NOVA results using three head-training seeds and 32 pairs per category on both datasets. NOVA values are seed means; Normal-only uses the same fixed normal-geometry model. Anomaly-ShapeNet uses the original 40 categories only.

Table 12: **Per-category detection results on Real3D-AD.** NOVA and Normal-only use the same normal-geometry model. Synthetic supervision improves all four category averages.

	Object-level				Point-level			
	AUROC		AP		AUROC		AP	
Category	Normal-only	NOVA	Normal-only	NOVA	Normal-only	NOVA	Normal-only	NOVA
Airplane	0.7088	0.8009	0.7334	0.8270	0.8757	0.9461	0.2282	0.3225
Candybar	0.9488	0.9700	0.9516	0.9755	0.9164	0.9620	0.2020	0.5450
Car	0.8552	0.8907	0.8670	0.9220	0.9439	0.9730	0.4370	0.6530
Chicken	0.8504	0.8420	0.8868	0.8369	0.8541	0.8993	0.3825	0.4789
Diamond	0.9832	0.9893	0.9827	0.9906	0.9285	0.9763	0.4903	0.7024
Duck	0.7588	0.8403	0.6695	0.7310	0.9146	0.9637	0.0602	0.0961
Fish	1.0000	1.0000	1.0000	1.0000	0.9780	0.9940	0.8463	0.9048
Gemstone	0.7260	0.7967	0.6353	0.6899	0.8693	0.9071	0.0404	0.0542
Seahorse	0.9976	0.9972	0.9977	0.9974	0.9657	0.9708	0.7994	0.8064
Shell	0.9295	0.9370	0.8317	0.8638	0.9827	0.9869	0.7313	0.7311
Starfish	0.9616	0.9689	0.9658	0.9698	0.9720	0.9780	0.7733	0.7946
Toffees	0.9044	0.9203	0.9230	0.9472	0.9209	0.9616	0.1191	0.1726
Macro	0.8854	0.9128	0.8704	0.8959	0.9268	0.9599	0.4258	0.5218

Table 13: **Per-category detection results on Anomaly-ShapeNet.** NOVA and Normal-only are evaluated on the same 40 categories used for training. Synthetic supervision improves all four category averages.

	Object-level				Point-level			
	AUROC		AP		AUROC		AP	
Category	Normal-only	NOVA	Normal-only	NOVA	Normal-only	NOVA	Normal-only	NOVA
ashtray0	1.0000	1.0000	1.0000	1.0000	0.9777	0.9802	0.7526	0.7603
bag0	0.9810	0.9810	0.9805	0.9805	0.9835	0.9876	0.6709	0.6903
bottle0	1.0000	1.0000	1.0000	1.0000	0.9946	0.9827	0.9091	0.9264
bottle1	1.0000	1.0000	1.0000	1.0000	0.9325	0.9405	0.5248	0.5361
bottle3	1.0000	1.0000	1.0000	1.0000	0.9604	0.9628	0.5662	0.5726
bowl0	1.0000	1.0000	1.0000	1.0000	0.9916	0.9942	0.9143	0.9342
bowl1	0.9370	0.9852	0.9619	0.9890	0.9662	0.9666	0.7599	0.8476
bowl2	1.0000	1.0000	1.0000	1.0000	0.9657	0.9836	0.7783	0.8596
bowl3	1.0000	1.0000	1.0000	1.0000	0.9911	0.9906	0.8738	0.8684
bowl4	0.9926	0.9827	0.9940	0.9871	0.9890	0.9879	0.8796	0.8934
bowl5	1.0000	1.0000	1.0000	1.0000	0.9898	0.9967	0.8864	0.9273
bucket0	0.9746	0.9735	0.9869	0.9865	0.8802	0.8901	0.2861	0.2744
bucket1	0.9079	0.9397	0.9509	0.9649	0.9708	0.9765	0.7823	0.8345
cap0	0.8593	0.8667	0.9218	0.9238	0.9775	0.9774	0.9171	0.9256
cap3	0.9053	0.8982	0.9471	0.9432	0.9709	0.9761	0.7843	0.8600
cap4	0.9228	0.9123	0.9528	0.9498	0.9819	0.9637	0.6228	0.7312
cap5	0.9895	0.9848	0.9917	0.9885	0.9954	0.9967	0.6961	0.7576
cup0	1.0000	1.0000	1.0000	1.0000	0.9966	0.9980	0.8670	0.8801
cup1	1.0000	1.0000	1.0000	1.0000	0.9814	0.9868	0.8274	0.8472
eraser0	1.0000	1.0000	1.0000	1.0000	0.9853	0.9831	0.8755	0.9120
headset0	1.0000	1.0000	1.0000	1.0000	0.9693	0.9764	0.3705	0.4025
headset1	1.0000	1.0000	1.0000	1.0000	0.9698	0.9748	0.5038	0.5146
helmet0	0.8696	0.8667	0.9294	0.9358	0.9449	0.9545	0.7183	0.7639
helmet1	1.0000	1.0000	1.0000	1.0000	0.9886	0.9919	0.8464	0.8585
helmet2	0.8899	0.8251	0.9466	0.9169	0.9538	0.9420	0.7318	0.7674
helmet3	0.9879	0.9919	0.9923	0.9945	0.9975	0.9984	0.9242	0.9331
jar0	1.0000	1.0000	1.0000	1.0000	0.9317	0.9312	0.2828	0.2760
microphone0	1.0000	1.0000	1.0000	1.0000	0.9758	0.9821	0.8456	0.9015
shelf0	0.8522	0.8300	0.9278	0.9110	0.9129	0.8975	0.6695	0.6976
tap0	1.0000	0.9525	1.0000	0.9711	0.9868	0.9861	0.8329	0.8652
tap1	0.9852	0.9617	0.9881	0.9712	0.9912	0.9834	0.6924	0.6778
vase0	1.0000	1.0000	1.0000	1.0000	0.9896	0.9913	0.8514	0.8777
vase1	0.9952	0.9968	0.9952	0.9968	0.8559	0.8756	0.6409	0.6579
vase2	1.0000	1.0000	1.0000	1.0000	0.9731	0.9795	0.3298	0.3232
vase3	0.8606	0.8636	0.9291	0.9308	0.9723	0.9711	0.8195	0.8524
vase4	0.8909	0.9455	0.9506	0.9721	0.9839	0.9855	0.8497	0.9170
vase5	1.0000	1.0000	1.0000	1.0000	0.9773	0.9935	0.8670	0.9247
vase7	1.0000	1.0000	1.0000	1.0000	0.9961	0.9985	0.9046	0.9194
vase8	0.7909	0.8556	0.8997	0.9285	0.9668	0.9684	0.8377	0.8749
vase9	0.8545	0.9333	0.9333	0.9589	0.9561	0.9589	0.7980	0.8826
Macro	0.9612	0.9637	0.9795	0.9800	0.9694	0.9716	0.7373	0.7682

E LIMITATIONS

NOVA derives deformation patterns from aligned normal objects of the same category, so the range of synthesized shapes depends on the variation present in the normal training set; categories whose normal objects are nearly identical provide fewer distinct patterns. The test-time registration diagnostic in Appendix D.2 evaluates the main detector and shows modest dataset-average changes alongside substantial localization losses on some categories. Sensitivity of synthesis to misaligned normal surfaces remains to be established. The evaluation covers geometric anomalies in two public point-cloud benchmarks without color or texture, and the relative effectiveness of synthesis strategies differs between them.