AMPROP3D: ADAPTIVE MATCHING WITH PROBABILITY PROPAGATION FOR 3D ANOMALY DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Feature matching-based methods have achieved promising performance in industrial point cloud anomaly detection, yet even the same defect type can exhibit sample-dependent feature responses, while different defect patterns activate distinct latent representations and exhibit spatial continuity. Existing methods struggle to capture these variations through sample-adaptive feature matching and neighborhood-aware anomaly scoring, limiting their ability to localize diverse defects. To address this issue, we propose a sample-adaptive point cloud anomaly detection method that jointly enhances sample-relevant feature responses and local consistency of defect responses among neighboring regions. Specifically, Adaptive Kernel Distribution Matching (AKDM) estimates channel relevance from the distribution discrepancies between local test regions and normal references, and incorporates the relevance into an adaptive kernel. Maximum mean discrepancy is then employed to measure distribution-level deviations, enabling feature matching to selectively emphasize latent responses relevant to the current sample. Building on this, Anomaly Probability Surface Propagation (APSP) propagates anomaly evidence among spatially proximal and feature-consistent regions, reinforcing contiguous defect responses while suppressing isolated noise. Experiments on MiniShiftAD, Anomaly-ShapeNet, Real3D-AD, and MulSen-AD demonstrate that the proposed method improves anomaly localization across multiple competitive 3D anomaly detection baselines, validating the effectiveness of sample-adaptive feature matching and spatial consistency modeling.

1 INTRODUCTION

Industrial 3D anomaly detection aims to identify defective products and accurately localize anomalous regions in point clouds, and plays an important role in intelligent manufacturing and automated quality inspection. Since real-world defects are typically rare, diverse, and costly to collect, unsupervised anomaly detection has become a practical setting, where only normal samples are available for modeling normal geometry and deviations from it are regarded as anomalies Bergmann et al. (2022); Liu et al. (2023). Compared with 2D images, point clouds directly encode local surface geometry and spatial structure, providing distinctive cues for detecting structural deformation, missing components, surface damage, and subtle geometric irregularities. Recent benchmarks such as Real3D-AD, Anomaly-ShapeNet, and MiniShiftAD further highlight the importance of fine-grained geometric anomaly localization, where defects may occupy only small surface regions Liu et al. (2023); Cheng et al. (2026b); Zhu et al. (2024). Existing approaches mainly follow reconstruction- and feature matching-based paradigms. Reconstruction methods learn normal geometric priors and detect anomalies from reconstruction errors Zhou et al. (2024); Cheng et al. (2025a); Ye et al. (2025); Zavrtanik et al. (2024), whereas feature matching methods compare query features against normal references stored in a memory bank Roth et al. (2022); Horwitz & Hoshen (2023). Owing to their simplicity and efficiency, feature matching approaches have been widely adopted, with recent methods improving performance through multimodal feature memories, geometric registration, and multi-scale local representations Wang et al. (2023); Liu et al. (2023); Cheng et al. (2026b).

Despite these advances, existing feature matching methods Cheng et al. (2026b); Roth et al. (2022); Horwitz & Hoshen (2023) do not fully account for the sample-dependent responses in latent feature space or the spatial continuity of industrial defects. Different defect patterns and samples can activate distinct latent feature responses, whereas conventional matching typically evaluates them

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.

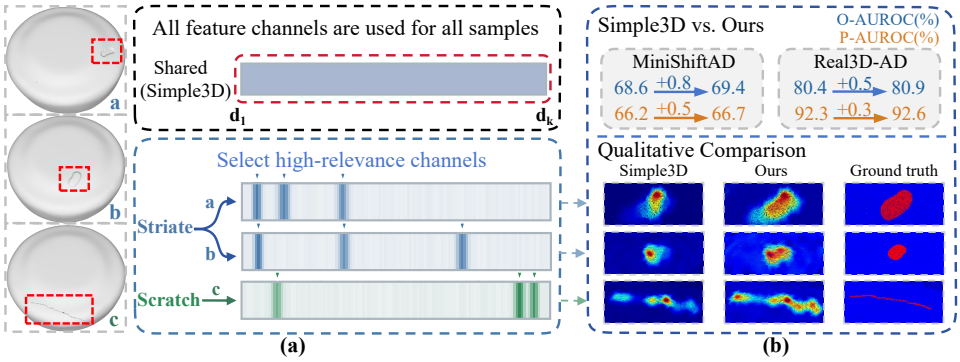

Figure 1: **Motivation for sample- and defect-adaptive feature weighting.** (a) Even for the same object and defect type, different samples exhibit distinct channel-wise relevance patterns, while different defects activate different relevant channels. A shared metric may therefore retain irrelevant responses, whereas adaptive weighting better emphasizes discriminative defect cues. (b) Sample- and defect-adaptive feature weighting improves quantitative and qualitative anomaly localization performance on two benchmarks compared with Simple3D Cheng et al. (2026b).

under a shared feature metric, allowing defect-irrelevant variations to interfere with discriminative anomaly cues. As illustrated in Fig. 1(a), even the same object with the same defect type can exhibit distinct relevance distributions across feature dimensions across different samples, while different defect types may also activate different latent responses. This observation suggests that anomaly matching should adaptively emphasize defect-relevant features for each sample, rather than relying on a shared feature metric. Fig. 1(b) further shows that such adaptive feature weighting yields more concentrated anomaly responses and improves quantitative localization performance. This issue is particularly important for subtle defects, whose abnormality may appear only as weak but consistent deviations along a limited subset of latent dimensions and can therefore be diluted by conventional point-wise matching. Meanwhile, most existing methods estimate anomaly scores independently for individual local regions, overlooking the fact that real defects are often spatially contiguous and locally consistent in geometric appearance. Weak anomalous responses therefore cannot benefit from supporting evidence in neighboring defective regions, while isolated false activations may remain prominent. Simple spatial smoothing cannot fully address this issue because it may propagate erroneous evidence across geometrically unrelated regions. These observations motivate two complementary questions: *which latent feature responses should be emphasized for each defect instance*, and *which neighboring anomaly evidence should reinforce each other*.

To address these issues, we propose a sample-adaptive point cloud anomaly detection framework that jointly models sample-relevant feature responses and spatially coherent anomaly evidence. Rather than directly comparing individual descriptors under a shared feature metric, we organize a local test neighborhood and its query-conditioned normal references as two feature distributions. Their discrepancy is used to adaptively determine the relevance of latent feature dimensions, allowing anomaly comparison to focus on responses that are more informative for the current query. We further represent local anomaly scores as a probability surface over the point cloud and propagate anomaly evidence among regions that are both spatially proximal and feature-consistent. In this way, the proposed framework adapts feature matching to heterogeneous defect patterns while exploiting local structural consistency to obtain more coherent and reliable anomaly localization.

Specifically, we first extract local geometric descriptors and construct multi-scale region-level representations, while features from normal training samples are stored in a normal memory bank. For each test region, we construct a local test set and retrieve a query-conditioned normal reference set for distribution-level comparison. We then introduce Adaptive Kernel Distribution Matching (AKDM), which estimates dimension-wise relevance from the distribution discrepancy between the two local sets and incorporates the resulting relevance into an adaptive Automatic Relevance Determination (ARD) kernel. Maximum Mean Discrepancy (MMD) is subsequently employed to quantify their distribution shift, producing sample-adaptive anomaly responses that selectively emphasize informative latent dimensions. Building on these responses, we further propose Anomaly Probability Surface Propagation (APSP), which maps local anomaly scores onto a probability surface and propagates anomaly evidence according to both spatial proximity and feature consistency. APSP strengthens coherent defect regions, recovers weak responses supported by structurally consistent

neighbors, and suppresses isolated activations lacking neighborhood support. Extensive experiments on MiniShiftAD, Anomaly-ShapeNet, Real3D-AD, and MulSen-AD demonstrate consistent improvements over competitive feature matching-based baselines, validating the effectiveness of sample-adaptive distribution matching and structure-aware anomaly propagation.

The main contributions of this work are summarized as follows:

- We develop a sample-adaptive feature matching framework for 3D anomaly detection that extends conventional point-wise matching to local distribution comparison and further incorporates structure-aware anomaly refinement, enabling adaptation to both sample-dependent latent responses and spatially coherent anomaly patterns.
- We introduce two complementary designs within this framework: Adaptive Kernel Distribution Matching (AKDM), which estimates dimension-wise relevance from local test–reference distribution discrepancies and incorporates it into an adaptive ARD kernel with MMD, and Anomaly Probability Surface Propagation (APSP), which propagates anomaly evidence among spatially proximal and feature-consistent regions to reinforce coherent defect responses while suppressing isolated activations.
- Extensive experiments on MiniShiftAD, Anomaly-ShapeNet, Real3D-AD, and MulSen-AD demonstrate consistent improvements over competitive feature matching-based methods, with particularly strong gains in point-level anomaly localization, validating the effectiveness and complementarity of the proposed components.

2 Related Work

3D Anomaly Detection. 3D anomaly detection has attracted growing attention for its ability to capture geometric information from point clouds Bergmann et al. (2022); Liu et al. (2023); Li et al. (2024). Existing methods mainly follow reconstruction-based Zhou et al. (2024); Cheng et al. (2025a); Ye et al. (2025); Zha et al. (2026) and feature-matching paradigms Roth et al. (2022); Horwitz & Hoshen (2023); Cao et al. (2024a); Cheng et al. (2025b). The former detects defects via reconstruction errors but may over-reconstruct subtle anomalies, while the latter compares test features with normal references for anomaly scoring. Recent feature-based methods exploit multimodal representations Wang et al. (2023); Chen et al. (2023); Long et al. (2025), registration-aware matching Liu et al. (2023); Yu et al. (2025), high-resolution descriptors Bergmann & Sattlegger (2023); Cheng et al. (2026b), and spatial aggregation Liang et al. (2025). Transferable point-cloud representations Zhao et al. (2021); Yu et al. (2022); Pang et al. (2022) and zero-shot vision-language approaches Cao et al. (2024b); Cheng et al. (2026a) have also been explored. Nevertheless, feature-matching methods often rely on fixed metrics and limited spatial interaction, which may obscure sample-specific cues and produce fragmented anomaly responses.

Distribution Matching and Structure-Aware Refinement. Maximum mean discrepancy (MMD) measures distribution discrepancies through kernel mean embeddings and is widely used for two-sample testing and distribution alignment Gretton et al. (2012). ARD kernels and metric learning further allow feature dimensions to receive different importance Rasmussen & Williams (2005); Weinberger & Saul (2009). However, most 3D feature-matching methods still rely on fixed nearest-neighbor metrics, leaving locally adaptive distribution matching underexplored. Structure-aware refinement has been studied in label propagation, manifold regularization, non-local filtering, and anisotropic diffusion Zhou et al. (2003); Belkin et al. (2006); Buades et al. (2005); Perona & Malik (1990). However, spatial-only smoothing may spread noisy responses across feature-inconsistent regions. In contrast, our method combines ARD-guided MMD with anomaly probability surface propagation using both spatial proximity and feature consistency.

Adaptive Feature Modeling for Anomaly Detection. Recent anomaly detection studies increasingly explore feature adaptation to better align pretrained representations with target domains. Existing approaches adapt patch descriptors, multimodal features, feature projections, prompts, prototypes, or structural granularity Lee et al. (2022); Tu et al. (2024); Gui et al. (2022); Cao et al. (2024b); Li et al. (2026); Kang et al. (2026). However, few explicitly construct a query-specific feature metric from local test–normal distribution discrepancies. In contrast, our AKDM transforms these discrepancies into channel-wise relevance weights for an adaptive ARD kernel, enabling sample-adaptive feature matching.

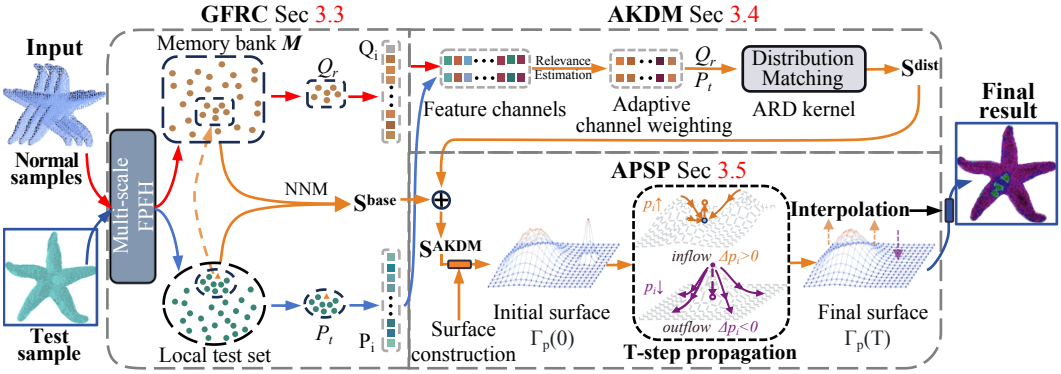

Figure 2: **Overview of the proposed framework.** Local test feature sets and their normal references are constructed from test features and a normal memory bank. Adaptive kernel distribution matching reweights feature channels and measures local distribution discrepancies via MMD. The resulting distribution scores are fused with the baseline scores to form an anomaly probability surface. Probability propagation then reinforces coherent anomaly responses in the joint spatial–feature space while suppressing isolated activations, yielding the final anomaly map.

3 Methodology

3.1 Problem Definition for 3D Anomaly Detection

3D anomaly detection follows a one-class setting, in which only normal samples are available during training. Let $\mathcal{D}_{\text{tr}} = \{X_i\}_{i=1}^{N}$ denote the normal training set, where $X_i = \{\mathbf{p}_{i,j}\}_{j=1}^{M_i}$ is a point cloud with M_i points and $\mathbf{p}_{i,j} \in \mathbb{R}^3$. Given a test point cloud $X^{\text{te}} = \{\mathbf{p}_j\}_{j=1}^{M}$, the detector determines whether it deviates from normal geometry and outputs an object-level anomaly score s^{obj} and a point-level anomaly map $\mathbf{S}^{\text{pt}} = \{s_j^{\text{pt}}\}_{j=1}^{M}$, where s_j^{pt} represents the anomaly score of point $\mathbf{p}_j$.

3.2 Overview

As illustrated in Fig. 2, our framework is built upon an FPFH-based feature-matching detector. Descriptors extracted from normal regions are stored in a memory bank, and nearest-neighbor matching provides the initial anomaly scores. For each test region i, Geometric Feature and Reference Construction (GFRC) constructs a spatial test set $\mathcal{P}_i$ and retrieves a query-conditioned normal reference set $\mathcal{Q}_i$. Based on these two sets, Adaptive Kernel Distribution Matching (AKDM) estimates channel-wise relevance and measures their distribution discrepancy using maximum mean discrepancy (MMD) with an ARD-guided kernel. The resulting distribution-aware anomaly score is then fused with the baseline score. Subsequently, APSP propagates the fused anomaly evidence in a joint spatial–feature space defined by local geometry and feature consistency, enhancing coherent responses while suppressing isolated noise. Finally, the region-level scores are interpolated back to the original point cloud to obtain the point-level anomaly map.

3.3 Geometric Feature and Reference Construction

For each point cloud X, we extract multi-scale Fast Point Feature Histogram (FPFH) descriptors and apply local feature spatial aggregation (Rusu et al., 2009; Cheng et al., 2026b). This produces $\mathcal{F}(X) = \{(\mathbf{f}_i, \mathbf{c}_i)\}_{i=1}^{n_X}$, where $\mathbf{f}_i \in \mathbb{R}^d$ represents the local geometry around spatial center $\mathbf{c}_i$. Descriptors from the normal training set are pooled into a memory bank $\mathcal{M} = \{\mathbf{m}_j\}_{j=1}^{K}$. We standardize each feature dimension using statistics estimated only from $\mathcal{M}$ and denote the standardized test and memory features by $\widehat{\mathbf{f}}_i$ and $\widehat{\mathbf{m}}_j$, respectively. The baseline anomaly score is defined as $s_i^{\text{base}} = \min_{\widehat{\mathbf{m}}_j \in \widehat{\mathcal{M}}} \|\widehat{\mathbf{f}}_i - \widehat{\mathbf{m}}_j\|_2$, namely the Euclidean distance from the test feature to its nearest standardized normal feature.

Although efficient, pointwise matching may overlook subtle deviations that are consistently shared across neighboring regions, since an individual descriptor can still remain close to its nearest normal

reference. To capture such local discrepancies, we replace isolated feature comparisons with two query-adaptive local sets:

$$\mathcal{P}_i = \left\{ \widehat{\mathbf{f}}_\ell \;\middle|\; \mathbf{c}_\ell \in \mathrm{kNN}_{k_l}\left(\mathbf{c}_i; \{\mathbf{c}_q\}_{q=1}^{n}\right) \right\}, \quad \mathcal{Q}_i = \mathrm{kNN}_{k_r}\left(\widehat{\mathbf{f}}_i; \widehat{\mathcal{M}}\right). \tag{1}$$

For each region i, the local test set $\mathcal{P}_i$ consists of the k_l standardized test features whose spatial centers are closest to $\mathbf{c}_i$, including the query itself, thereby characterizing its local spatial context. In parallel, the normal reference set $\mathcal{Q}_i$ is formed by the k_r standardized memory features nearest to $\widehat{\mathbf{f}}_i$ in feature space, providing normal anchors tailored to the current query. Since $\mathcal{Q}_i$ is selected according to $\widehat{\mathbf{f}}_i$, the discrepancy between $\mathcal{P}_i$ and $\mathcal{Q}_i$ reflects a query-adaptive deviation from the corresponding normal feature distribution. Based on these local sets, AKDM further measures the distribution discrepancy, as described in the next subsection.

3.4 Adaptive Kernel Distribution Matching

Given the local test set $\mathcal{P}_i$ and its query-conditioned normal reference set $\mathcal{Q}_i$, we measure their distribution-level discrepancy using maximum mean discrepancy (MMD) in the shared standardized feature space. MMD compares the mean embeddings of two distributions in a reproducing kernel Hilbert space (RKHS) and can be evaluated solely through kernel functions Gretton et al. (2012). We employ the biased empirical estimator, whose detailed formulation is provided in the Appendix.

A fixed kernel assigns equal importance to all feature channels, although different industrial defects may manifest through different geometric cues. Consequently, the most discriminative feature dimensions can vary across test regions, while irrelevant or highly variable channels may introduce noise into the similarity measure. For each local pair $(\mathcal{P}_i, \mathcal{Q}_i)$, we define the standardized mean shift and relative variation of channel r as:

$$\delta_{i,r} = \frac{(\mu_{i,r}^{P} - \mu_{i,r}^{Q})^2}{v_{i,r}^{Q} + \epsilon}, \qquad u_{i,r} = \frac{v_{i,r}^{P}}{v_{i,r}^{Q} + \epsilon}. \tag{2}$$

We then compute and normalize the channel relevance:

$$\rho_{i,r} = \frac{a + \frac{1}{2}\delta_{i,r}}{b + \omega u_{i,r} + \epsilon}, \qquad w_{i,r} = \frac{\exp(\rho_{i,r}/\tau)}{\sum_{\ell=1}^{d} \exp(\rho_{i,\ell}/\tau)}. \tag{3}$$

Here, μ and v denote the empirical mean and variance, respectively; a and b are stabilizing constants, ω controls the variation penalty, and τ is the softmax temperature. Thus, channels with larger distribution shifts and lower local variation receive higher weights.

The relevance weights are incorporated into an automatic relevance determination (ARD) kernel:

$$k_i^{\mathrm{ARD}}(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{1}{2\sigma^2} \sum_{r=1}^{d} w_{i,r}(x_r - y_r)^2 \right), \tag{4}$$

where σ denotes the kernel bandwidth. This adaptive kernel emphasizes channels carrying reliable anomaly-related deviations while suppressing those dominated by intrinsic variation.

Using k_i^{ARD}, we compute the distribution-level anomaly score as $s_i^{\mathrm{dist}} = \mathrm{MMD}(\mathcal{P}_i, \mathcal{Q}_i; k_i^{\mathrm{ARD}})$. After min–max normalization, the distribution score is fused with the feature-matching score:

$$s_i^{\mathrm{AKDM}} = s_i^{\mathrm{base}} + \lambda \widetilde{s}_i^{\mathrm{dist}}, \tag{5}$$

λ controls the contribution of the distribution-level correction. In this way, AKDM preserves the original nearest-neighbor evidence while complementing it with query-adaptive local distribution discrepancies. The resulting score s_i^{AKDM} is subsequently processed by APSP.

3.5 Anomaly Probability Surface Propagation

Although AKDM captures local distributional deviations, independently estimated region-level scores may still lead to fragmented responses over spatially continuous defects. A straightforward spatial smoothing strategy can improve response continuity, but may also propagate anomalous evidence across geometrically unrelated surfaces. To address this limitation, we introduce Anomaly

Probability Surface Propagation (APSP), which propagates anomaly evidence over a joint spatial–feature space defined by spatial proximity and feature consistency.

Given the AKDM scores $\{s_i^{\mathrm{AKDM}}\}_{i=1}^n$, we apply min–max normalization within each test point cloud to obtain the initial anomaly likelihoods $p_i^{(0)} \in [0,1]$. These values define a discrete anomaly probability surface over the local region centers $\{\mathbf{c}_i\}_{i=1}^n$. Here, "probability" refers to a normalized relative anomaly likelihood.

For each region i, we construct a neighborhood $\mathcal{N}_i$ containing its k_p nearest region centers, excluding i itself. The centers follow the same coordinate normalization as the input point cloud. We define a joint spatial–feature discrepancy between regions i and j, which characterizes their relation in the feature-aware propagation space:

$$D_i(j) = \frac{\|\mathbf{c}_i - \mathbf{c}_j\|_2^2}{\sigma_c^2} + \frac{1}{\sigma_f^2} \sum_{r=1}^{d} w_{i,r} \left(\widehat{f}_{i,r} - \widehat{f}_{j,r} \right)^2, \tag{6}$$

where $\sigma_c, \sigma_f > 0$ control the spatial and feature bandwidths. $\widehat{\mathbf{f}}_i$ and $\widehat{\mathbf{f}}_j$ are the standardized local features. The relevance weights $w_{i,r}$ are inherited from the AKDM comparison for region i and satisfy $w_{i,r} \geq 0$ and $\sum_r w_{i,r} = 1$.

We convert the distance into a conductivity coefficient and normalize it within $\mathcal{N}_i$:

$$\kappa_{ij} = \exp\left[-\frac{1}{2} D_i(j) \right], \quad a_{ij} = \frac{\kappa_{ij}}{\sum_{\ell \in \mathcal{N}_i} \kappa_{i\ell}}, \qquad j \in \mathcal{N}_i. \tag{7}$$

A larger a_{ij} indicates that region j is more suitable for propagating evidence to region i. The resulting conductivities are locally adaptive and asymmetric, reflecting query-specific relevance weights and neighborhood structure. Their normalization satisfies $\sum_{j \in \mathcal{N}_i} a_{ij} = 1$, ensuring normalized evidence aggregation during probability-surface evolution.

At iteration t, the directed anomaly-probability flow from region j to region i is defined as:

$$\Phi_{j \to i}^{(t)} = a_{ij} \left(p_j^{(t)} - p_i^{(t)} \right). \tag{8}$$

The flow characterizes the local evolution tendency induced by neighboring anomaly evidence. A region receives positive flow when a structurally consistent neighbor has a higher anomaly likelihood. Conversely, an isolated response tends to be attenuated when its weighted neighborhood exhibits lower anomaly likelihoods. We evolve the anomaly-probability surface by combining the structural flow with a fidelity term:

$$p_i^{(t+1)} = p_i^{(t)} + \alpha \sum_{j \in \mathcal{N}_i} \Phi_{j \to i}^{(t)} + (1 - \alpha) \left(p_i^{(0)} - p_i^{(t)} \right), \tag{9}$$

where $0 \leq \alpha < 1$ balances structure-aware probability flow against the retention of the initial anomaly evidence. The fidelity term retains the AKDM response, while the flow term encourages consistency among spatially and feature-compatible neighbors. With fixed, normalized conductivities and $0 \leq \alpha < 1$, the evolution remains bounded and converges to a unique equilibrium.

After T iterations, we use the change in anomaly likelihood to correct the original AKDM score:

$$s_i^{\mathrm{APSP}} = s_i^{\mathrm{AKDM}} \left[1 + \beta \left(p_i^{(T)} - p_i^{(0)} \right) \right], \qquad 0 \leq \beta \leq 1. \tag{10}$$

The residual correction strengthens regions supported by consistent neighbors and suppresses isolated responses, while preserving score non-negativity for $0 \leq \beta \leq 1$.

APSP produces a processed anomaly score s_i^{APSP} for each FPS-sampled region center $\mathbf{c}_i$. To obtain a dense point-wise anomaly map, we interpolate the regional scores to every input point $\mathbf{x}_q$ using its three nearest region centers:

$$s_q^{\mathrm{point}} = \sum_{i \in \mathcal{N}_3(\mathbf{x}_q)} \omega_{qi} s_i^{\mathrm{APSP}}, \qquad \omega_{qi} = \frac{(\|\mathbf{x}_q - \mathbf{c}_i\|_2^2 + \epsilon_I)^{-1}}{\sum_{j \in \mathcal{N}_3(\mathbf{x}_q)} (\|\mathbf{x}_q - \mathbf{c}_j\|_2^2 + \epsilon_I)^{-1}}, \tag{11}$$

Table 1: Comparison of anomaly detection performance on four 3D anomaly detection benchmarks.

Method →	PatchCore-FP	PatchCore-PM	PatchCore-PB	R3D-AD	GLFM	Reg2Inv	Simple3D	**AMProp3D**
MiniShiftAD	CVPR'22	CVPR'22	CVPR'22	ECCV'24	TASE'25	NeurIPS'25	AAAI'26	**Ours**
O-AUROC/P-AUROC	65.1/53.7	55.9/54.1	50.0/51.4	53.6/49.8	55.8/58.7	61.1/67.6	68.6/66.2	**70.0/69.8**
Method →	M3DM-PB	PatchCore-FP	PatchCore-PM	IMRNet	Reg3D-AD	GLFM	Simple3D	**AMProp3D**
MulSen-AD	CVPR'23	CVPR'22	CVPR'22	CVPR'24	NeurIPS'23	TASE'25	AAAI'26	**Ours**
O-AUROC/P-AUROC	70.5/61.1	86.0/64.0	84.0/60.5	60.1/46.7	74.9/64.1	78.5/66.5	88.2/80.3	**90.5/83.2**
Method →	PatchCore	Reg3D-AD	IMRNet	PO3AD	MC3D-AD	HPPF	Simple3D	**AMProp3D**
Anomaly-ShapeNet	CVPR'22	NeurIPS'23	CVPR'24	CVPR'25	IJCAI'25	CVPR'26	AAAI'26	**Ours**
O-AUROC/P-AUROC	56.8/58.0	57.2/66.8	66.1/65.0	83.9/89.8	84.2/74.8	87.6/–	86.0/92.9	**88.5/93.9**
Method →	PatchCore	Reg3D-AD	IMRNet	PO3AD	MC3D-AD	GLFM	Simple3D	**AMProp3D**
Real3D-AD	CVPR'22	NeurIPS'23	CVPR'24	CVPR'25	IJCAI'25	TASE'25	AAAI'26	**Ours**
O-AUROC/P-AUROC	68.2/69.2	70.4/70.0	72.5/–	76.5/–	78.2/76.8	75.0/76.7	80.4/92.3	**81.2/93.0**

Table 2: **Category-level quantitative results on MiniShiftAD.** Results are reported as O-AUROC (%)/P-AUROC (%). The best result is shown in **bold**, and the second-best result is underlined.

Method →	**PatchCore-FP**	**PatchCore-PM**	**PatchCore-PB**	**R3D-AD**	**GLFM**	**Reg2Inv**	**Simple3D**	**AMProp3D**
Category ↓	**CVPR'22**	**CVPR'22**	**CVPR'22**	**ECCV'24**	**TASE'25**	**NeurIPS'25**	**AAAI'26**	**Ours**
Capsule	88.0/58.4	32.4/54.6	27.2/51.1	60.0/51.3	41.2/69.4	68.0/78.6	**92.9**/80.6	86.6/**86.4**
Cube	61.9/52.7	37.5/51.2	44.1/47.7	53.5/53.7	48.8/65.5	68.0/**77.1**	61.9/66.6	**78.5**/70.9
Spring pad	68.9/53.4	53.2/55.0	41.0/52.8	32.9/50.3	53.4/54.5	36.4/56.0	68.7/63.1	**77.3/63.7**
Screw	25.0/56.0	59.2/54.9	67.2/51.3	63.9/48.7	**68.2**/53.4	37.6/63.2	23.2/64.6	19.4/**71.1**
Screen	70.2/53.3	63.5/52.8	54.0/50.2	42.9/47.2	63.9/66.8	**88.3/83.2**	70.0/57.5	54.9/59.5
Piggy	61.3/55.5	58.4/53.8	50.3/52.4	36.0/49.7	63.8/53.4	**76.4/79.2**	67.4/67.5	68.5/72.3
Nut	69.2/54.5	58.0/55.1	43.2/53.0	66.3/53.7	49.4/56.6	70.8/63.5	76.9/73.1	**78.2/76.6**
Flat pad	60.7/54.1	59.9/57.8	60.9/52.7	59.3/49.5	67.7/57.6	28.2/35.7	68.8/61.8	**74.0/64.2**
Plastic cylinder	81.2/54.8	77.7/53.0	60.3/51.6	70.2/50.7	63.5/58.9	76.8/**73.7**	**83.2**/70.3	77.1/73.0
Button cell	86.2/57.4	64.3/52.7	51.1/48.4	43.7/50.3	72.0/62.3	71.1/77.7	**90.0**/73.2	86.8/**80.2**
Toothbrush	48.8/46.7	55.1/**56.7**	49.0/55.2	66.3/39.2	44.1/47.8	38.3/46.7	66.1/45.5	**74.2**/47.3
Light	59.8/47.7	51.2/51.6	51.3/50.5	48.2/53.0	33.0/58.2	**73.4/77.1**	54.2/70.1	64.9/72.6
Mean	65.1/53.7	55.9/54.1	50.0/51.4	53.6/49.8	55.8/58.7	61.1/67.6	68.6/66.2	**70.0/69.8**

where $\mathcal{N}_3(\mathbf{x}_q)$ contains the three nearest region centers and ϵ_I ensures numerical stability. The interpolated scores form the final point-level anomaly map S^{pt}. Let $\mathcal{T}_{K_o}$ denote the indices of its K_o highest scores. We compute the object-level anomaly score as $s^{\text{obj}} = K_o^{-1} \sum_{q \in \mathcal{T}_{K_o}} s_q^{\text{pt}}$. This aggregation retains the contribution of localized defects that could be diluted by averaging over the entire point cloud.

4 Experiments

4.1 Experimental Setup

Datasets. We extensively evaluate our method on four representative 3D anomaly detection benchmarks: MiniShiftAD Cheng et al. (2026b), Anomaly-ShapeNet Li et al. (2024), Real3D-AD Liu et al. (2023), and MulSen-AD Li et al. (2025).

Baselines. We compare our method with two broad families of representative 3D anomaly detection approaches: (1) **reconstruction- and regression-based methods**, including IMRNet Li et al. (2024), R3D-AD Zhou et al. (2024), MC3D-AD Cheng et al. (2025a), Reg2Inv Yu et al. (2025), and PO3AD Ye et al. (2025); and (2) **feature matching- and representation-based methods**, including PatchCore Roth et al. (2022), CPMF Cao et al. (2024a), M3DM Wang et al. (2023), Reg3D-AD Liu et al. (2023), GLFM Cheng et al. (2025b), ISMP Liang et al. (2025), PASDF Zheng et al. (2025), HPPF Kang et al. (2026), and Simple3D Cheng et al. (2026b). The reported results are obtained from the original publications or reproduced using publicly available implementations.

Metrics. To evaluate the effectiveness of our method, we employ the widely used Area Under the Receiver Operating Characteristic curve (AUROC) metric. Specifically, object-wise AUROC (O-AUROC) measures object-level anomaly detection performance, while point-wise AUROC (P-AUROC) evaluates point-level anomaly localization performance across all four datasets.

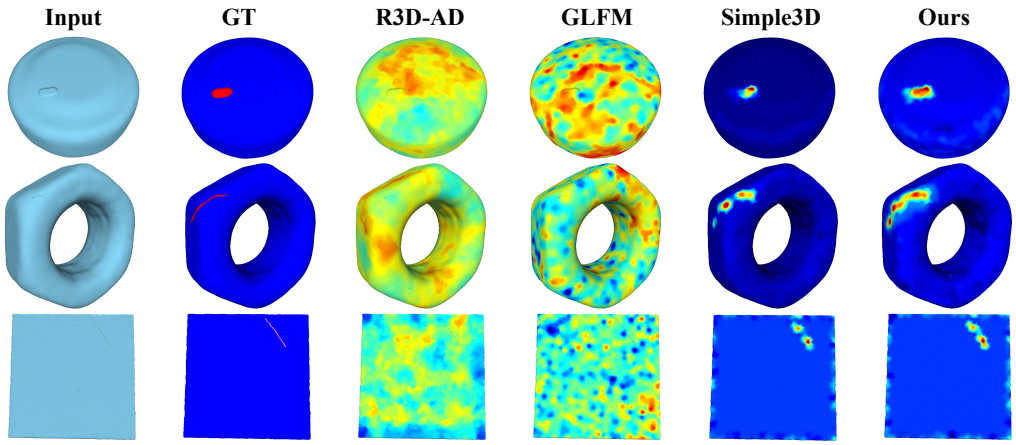

Figure 3: **Qualitative comparison on MiniShiftAD.** Compared with R3D-AD, GLFM, and Simple3D, our method yields more accurate and spatially concentrated anomaly responses.

4.2 Main Results

As shown in Table 1, our method achieves the best O-AUROC and P-AUROC across all four benchmarks, demonstrating consistent performance under high-resolution, multi-sensor, synthetic, and real-scanned settings. On MiniShiftAD and MulSen-AD, our method achieves O-AUROC/P-AUROC scores of 70.0%/69.8% and 90.5%/83.2%, outperforming the next-best results by 1.4/2.2 and 2.3/2.9 percentage points, respectively. On Anomaly-ShapeNet and Real3D-AD, it obtains 88.5%/93.9% and 81.2%/93.0%, with corresponding gains of 0.9/1.0 and 0.8/0.7 percentage points.

The category-wise results further demonstrate robustness across diverse object geometries, anomaly types, and sensing conditions. As shown in Table 2, our method achieves the highest P-AUROC in 6 of 12 MiniShiftAD categories. Moreover, 7 MulSen-AD categories achieve at least 93.8% P-AUROC, while P-AUROC exceeds 90% in 34 of 40 Anomaly-ShapeNet categories and 9 of 12 Real3D-AD categories.

4.3 Qualitative Results and Analysis

Fig. 3 presents qualitative comparisons of point-level anomaly localization on MiniShiftAD. Our method yields more concentrated responses within defect regions, better preserves their spatial extent, and suppresses scattered activations on normal surfaces. The advantage is particularly clear for subtle local defects, where competing methods often produce incomplete or diffuse responses. This behavior is consistent with AKDM emphasizing sample-relevant feature discrepancies and APSP reinforcing spatially coherent anomaly evidence. Overall, the results demonstrate more precise and coherent localization of fine-grained defects.

4.4 Ablation Study

To isolate the contributions of AKDM and APSP, we evaluate the baseline, each module individually, and their combination across all four datasets, with O-AUROC/P-AUROC reported in Table 3. Additional ablations are conducted on MulSen-AD, whose realistic scanning noise provides a representative testbed for these design choices.

Table 3: Ablation results of AKDM and APSP.

AKDM	✗	✓	✗	✓
APSP	✗	✗	✓	✓
MiniShiftAD	68.6/66.2	69.4/66.7	68.8/67.2	70.0/69.8
Anomaly-ShapeNet	86.0/92.9	87.3/93.2	87.8/93.5	88.5/93.9
Real3D-AD	80.4/92.3	80.9/92.6	80.7/92.8	81.2/93.0
MulSen-AD	88.2/80.3	89.1/82.5	89.4/82.1	90.5/83.2

Validation on main components. Table 3 evaluates the individual and combined contributions of AKDM and APSP. On average, AKDM alone improves O-AUROC and P-AUROC by 0.88 and 0.83 percentage points, respectively, while APSP alone provides gains of 0.88 and 0.98 points. Combin-

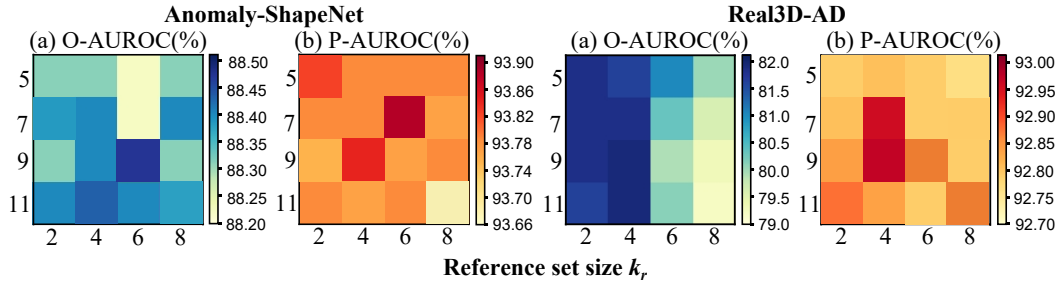

Figure 4: Sensitivity to the local neighborhood size k_l and reference set size k_r on Anomaly-ShapeNet and Real3D-AD. The vertical axis denotes k_l, while the horizontal axis denotes k_r. Each heatmap reports O-AUROC or P-AUROC under different parameter combinations.

ing the two modules achieves the best performance on all four datasets, with average improvements of 1.75 points in O-AUROC and 2.05 points in P-AUROC over the baseline. These gains demonstrate their complementary roles: AKDM enhances anomaly-aware feature matching, while APSP promotes coherent responses among related points. The larger P-AUROC gains on MiniShiftAD and MulSen-AD further highlight their benefit for point-level anomaly localization.

Validation on adaptive distribution matching. With APSP disabled, we compared AKDM against direct 1-NN, adaptive kNN, and local MMD on MulSen-AD. AKDM achieved 89.1% O-AUROC and 82.5% P-AUROC, surpassing the strongest competing strategy by 0.6 and 0.5 percentage points, respectively. The consistent gains on both metrics indicate that modeling local test–normal distribution discrepancies is more effective than relying on point-wise nearest-neighbor matching or fixed local distribution comparison. By further adapting the kernel to query-specific channel relevance, AKDM provides a more discriminative matching metric for anomaly detection and localization.

Validation on Bayesian-inspired ARD. To isolate the effect of feature-channel weighting, we compared Bayesian-inspired ARD with uniform weighting and mean-difference ARD under the same local MMD setting, with APSP disabled. Bayesian-inspired ARD achieved 89.1% O-AUROC and 82.5% P-AUROC, outperforming uniform weighting by 1.2/2.0 points and mean-difference ARD by 0.9/0.7 points. These gains indicate that incorporating uncertainty provides more reliable channel-relevance estimates and better emphasizes defect-discriminative features for anomaly localization.

Validation on spatial–feature propagation. Under the same AKDM configuration, we compared spatial-feature propagation with no propagation, spatial-only propagation, and feature-only propagation. The spatial-only and feature-only variants achieved 90.2%/82.8% and 89.5%/83.0% O-AUROC/P-AUROC, respectively, while combining both affinities achieved the best results of 90.5%/83.2%, improving over no propagation by 1.4/0.7 points. These results confirm the complementarity of spatial and feature affinities.

Hyperparameter sensitivity. We evaluated k_l and k_r on Anomaly-ShapeNet and Real3D-AD while keeping all other settings fixed. Performance remained stable across most configurations, although O-AUROC on Real3D-AD was more sensitive to k_r. This consistency suggests that the method is not overly sensitive to moderate changes in neighborhood or reference-set size, indicating reasonable robustness to these local sampling choices. While $k_r = 4$ performs favorably on these two datasets, the optimal setting varies across datasets. We therefore use a fixed configuration of $k_l = 7$ and $k_r = 8$ for all experiments to avoid dataset-specific tuning.

5 Conclusion

We presented a sample-adaptive framework for unsupervised 3D anomaly detection that integrates adaptive distribution matching with structure-aware probability propagation. AKDM adapts the matching metric to query-conditioned local distribution discrepancies, selectively emphasizing latent responses that are informative for the current defect pattern, while APSP refines anomaly evidence through propagation among spatially proximal and feature-consistent regions. Experiments on four benchmarks demonstrate the complementary effects of the two components and consistent improvements in anomaly localization, showing the effectiveness of sample-adaptive feature matching and spatially coherent evidence propagation for heterogeneous 3D defects.

AI USE STATEMENT

Generative AI tools were used to improve the clarity, grammar, concision, and readability of author-written text and to assist with routine code debugging, such as interpreting error messages and suggesting possible corrections. All AI-assisted textual revisions were reviewed and edited by the authors, and all code suggestions were manually inspected and validated through execution and testing before being incorporated.

Generative AI tools were not used to formulate or refine the research problem or hypotheses, develop the theoretical or conceptual framework, derive mathematical claims or proofs, design the proposed methodology or experiments, implement the proposed method, generate, clean, or preprocess data, perform qualitative data analysis, or interpret experimental results. The authors take full responsibility for the final manuscript, code, scientific claims, and conclusions.

ETHICS STATEMENT

All datasets used in this study are publicly available benchmarks for 3D anomaly detection and were employed solely for research and evaluation. This work does not involve human participants, personally identifiable information, sensitive personal attributes, or privately collected data. The proposed method is evaluated exclusively in an offline benchmarking setting and is not used to make decisions about individuals. To the best of our knowledge, this study does not raise specific ethical concerns regarding human-subject research, privacy, discrimination, or security.

REPRODUCIBILITY STATEMENT

The main paper describes the proposed method, algorithmic framework, datasets, evaluation metrics, and experimental protocol. Additional implementation details, hyperparameter settings, data preprocessing procedures, and extended experimental results are provided in the appendix and supplementary materials. The complete pseudocode of the proposed framework is also provided in the appendix to facilitate implementation and reproduction. All experiments are conducted on publicly available datasets. To facilitate independent reproduction of the reported results, we will publicly release the source code, configuration files, and experimental scripts.

REFERENCES

Mikhail Belkin, Partha Niyogi, and Vikas Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research*, 7(85):2399–2434, 2006. URL http://jmlr.org/papers/v7/belkin06a.html.

Paul Bergmann and David Sattlegger. Anomaly detection in 3d point clouds using deep geometric descriptors. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2613–2623, 2023. doi: 10.1109/WACV56688.2023.00264.

Paul Bergmann, Xin Jin, David Sattlegger, and Carsten Steger. The mvtec 3d-ad dataset for unsupervised 3d anomaly detection and localization. In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 202–213. SCITEPRESS - Science and Technology Publications, 2022. doi: 10.5220/0010865000003124. URL http://dx.doi.org/10.5220/0010865000003124.

A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pp. 60–65 vol. 2, 2005. doi: 10.1109/CVPR.2005.38.

Yunkang Cao, Xiaohao Xu, and Weiming Shen. Complementary pseudo multimodal feature for point cloud anomaly detection. *Pattern Recognition*, 156:110761, 2024a. ISSN 0031-3203. doi: 10.1016/j.patcog.2024.110761. URL https://www.sciencedirect.com/science/article/pii/S0031320324005120.

Yunkang Cao, Jiangning Zhang, Luca Frittoli, Yuqi Cheng, Weiming Shen, and Giacomo Boracchi. Adaclip: Adapting clip with hybrid learnable prompts for zero-shot anomaly detection.

In *Computer Vision – ECCV 2024*, pp. 55–72. Springer Nature Switzerland, 2024b. ISBN 9783031727610. doi: 10.1007/978-3-031-72761-0_4. URL http://dx.doi.org/10.1007/978-3-031-72761-0_4.

Ruitao Chen, Guoyang Xie, Jiaqi Liu, Jinbao Wang, Ziqi Luo, Jinfan Wang, and Feng Zheng. Easynet: An easy network for 3d industrial anomaly detection. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, pp. 7038–7046, 2023.

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 Thirty-Fourth International Joint Conference on Artificial Intelligence*, IJCAI-2025, pp. 837–845. International Joint Conferences on Artificial Intelligence Organization, 2025a. doi: 10.24963/ijcai.2025/94. URL http://dx.doi.org/10.24963/ijcai.2025/94.

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. doi: 10.1109/TASE.2025.3544462.

Yuqi Cheng, Yunkang Cao, Guoyang Xie, Zhichao Lu, and Weiming Shen. Toward zero-shot point cloud anomaly detection: A multiview projection framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 56(3):1747–1760, March 2026a. ISSN 2168-2216. doi: 10.1109/TSMC.2025.3648581.

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, 2026b.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012. URL http://jmlr.org/papers/v13/gretton12a.html.

Xingtai Gui, Di Wu, Yang Chang, and Shicai Fan. Constrained adaptive projection with pretrained features for anomaly detection. In Lud De Raedt (ed.), *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pp. 2059–2065. International Joint Conferences on Artificial Intelligence Organization, 7 2022. doi: 10.24963/ijcai.2022/286. URL https://doi.org/10.24963/ijcai.2022/286. Main Track.

Eliahu Horwitz and Yedid Hoshen. Back to the feature: Classical 3d features are (almost) all you need for 3d anomaly detection. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 2968–2977, 2023. doi: 10.1109/CVPRW59228.2023.00298.

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, 2026. URL https://arxiv.org/abs/2604.03972.

Sungwook Lee, Seunghyun Lee, and Byung Cheol Song. Cfa: Coupled-hypersphere-based feature adaptation for target-oriented anomaly localization. *IEEE Access*, 10:78446–78454, 2022. doi: 10.1109/ACCESS.2022.3193699.

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 (CVPR)*, pp. 22207–22216, June 2024.

Wenqiao Li, Bozhong Zheng, Xiaohao Xu, Jinye Gan, Fading Lu, Xiang Li, Na Ni, Zheng Tian, Xiaonan Huang, Shenghua Gao, and Yingna Wu. Multi-sensor object anomaly detection: Unifying appearance, geometry, and internal properties. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9984–9993, 2025. doi: 10.1109/CVPR52734.2025.00933.

Yufei Li, Long Tian, Yuyang Dai, Wenchao Chen, Liang Bao, and Xiyang Liu. Fastref: Fast prototype refinement for few-shot industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 43040–43049, June 2026.

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, 2025. URL https://arxiv.org/abs/2412.13461.

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 *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2023.

Kaifang Long, Guoyang Xie, Lianbo Ma, Jiaqi Liu, and Zhichao Lu. Revisiting multimodal fusion for 3d anomaly detection from an architectural perspective. AAAI Press, 2025. ISBN 978-1-57735-897-8. doi: 10.1609/aaai.v39i12.33337. URL https://doi.org/10.1609/aaai.v39i12.33337.

Yatian Pang, Wenxiao Wang, Francis EH Tay, Wei Liu, Yonghong Tian, and Li Yuan. Masked autoencoders for point cloud self-supervised learning. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part II*, pp. 604–621. Springer, 2022.

P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990. doi: 10.1109/34.56205.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005. ISBN 9780262256834. doi: 10.7551/mitpress/3206.001.0001. URL https://doi.org/10.7551/mitpress/3206.001.0001.

Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14298–14308, 2022. doi: 10.1109/CVPR52688.2022.01392.

Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *2009 IEEE International Conference on Robotics and Automation*, pp. 3212–3217, 2009. doi: 10.1109/ROBOT.2009.5152473.

Yuanpeng Tu, Boshen Zhang, Liang Liu, Yuxi Li, Jiangning Zhang, Yabiao Wang, Chengjie Wang, and Cairong Zhao. Self-supervised feature adaptation for 3d industrial anomaly detection. In Ales Leonardis, Elisa Ricci, Stefan Roth, Olga Russakovsky, Torsten Sattler, and Gül Varol (eds.), *Computer Vision - ECCV 2024 - 18th European Conference, Milan, Italy, September 29-October 4, 2024, Proceedings, Part II*, volume 15060 of *Lecture Notes in Computer Science*, pp. 75–91. Springer, 2024. ISBN 978-3-031-72627-9. doi: 10.1007/978-3-031-72627-9_5. URL https://doi.org/10.1007/978-3-031-72627-9_5.

Yue Wang, Jinlong Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, and Chengjie Wang. Multimodal industrial anomaly detection via hybrid fusion. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8032–8041, 2023. doi: 10.1109/CVPR52729.2023.00776.

Kilian Q. Weinberger and Lawrence K. Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10(9):207–244, 2009. URL http://jmlr.org/papers/v10/weinberger09a.html.

Jianan Ye, Weiguang Zhao, Xi Yang, Guangliang Cheng, and Kaizhu Huang. Po3ad: Predicting point offsets toward better 3d point cloud anomaly detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025.

Xumin Yu, Lulu Tang, Yongming Rao, Tiejun Huang, Jie Zhou, and Jiwen Lu. Point-bert: Pre-training 3d point cloud transformers with masked point modeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 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. 2025. URL https://arxiv.org/abs/2510.16865.

Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. Cheating depth: Enhancing 3d surface anomaly detection via depth simulation. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2153–2161, 2024. doi: 10.1109/WACV57701.2024.00216.

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 Sven Koenig, Chad Jenkins, and Matthew E. Taylor (eds.), *Fortieth AAAI Conference on Artificial Intelligence, Thirty-Eighth Conference on Innovative Applications of Artificial Intelligence, Sixteenth Symposium on Educational Advances in Artificial Intelligence, AAAI 2026, Singapore, January 20-27, 2026*, pp. 12340–12348. AAAI Press, 2026. doi: 10.1609/AAAI.V40I15.38226. URL https://doi.org/10.1609/aaai.v40i15.38226.

Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip Torr, and Vladlen Koltun. Point transformer. 2021. URL https://arxiv.org/abs/2012.09164.

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 *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 27063–27072, 2025. doi: 10.1109/ICCV51701.2025.02512.

Dengyong Zhou, Olivier Bousquet, Thomas Lal, Jason Weston, and Bernhard Schölkopf. Learning with local and global consistency. In S. Thrun, L. Saul, and B. Schölkopf (eds.), *Advances in Neural Information Processing Systems*, volume 16. MIT Press, 2003. URL https://proceedings.neurips.cc/paper_files/paper/2003/file/87682805257e619d49b8e0dfdc14affa-Paper.pdf.

Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3D: A modern library for 3D data processing. *arXiv:1801.09847*, 2018.

Zheyuan Zhou, Le Wang, Naiyu Fang, Zili Wang, Lemiao Qiu, and Shuyou Zhang. R3d-ad: Reconstruction via diffusion for 3d anomaly detection. In *European Conference on Computer Vision (ECCV)*, 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 32nd ACM International Conference on Multimedia*, pp. 4680–4689, 2024. doi: 10.1145/3664647.3680919.

APPENDIX CONTENTS

A Additional Implementation Details **15**

A.1 Geometric Feature and Reference Construction 15

A.2 Adaptive Kernel Distribution Matching 15

A.3 Parameter Settings 17

A.4 Overall Algorithm 17

B Theoretical Properties of APSP **19**

B.1 Matrix Form of Probability-Surface Evolution 19

B.2 Boundedness of the Probability Surface 19

B.3 Convergence and Unique Equilibrium 19

B.4 Bounded Residual Score Correction 20

B.5 Interpretation 20

C Extended Category-wise Results **21**

C.1 Anomaly-ShapeNet 21

C.2 MulSen-AD 21

C.3 Real3D-AD 21

D Additional Ablation Studies **22**

D.1 Matching Strategy 22

D.2 Bayesian-inspired ARD 23

D.3 Spatial-feature Propagation 23

D.4 Accuracy–Efficiency Analysis of AKDM and APSP 24

E Additional Qualitative Results **26**

E.1 MiniShiftAD 27

E.2 MulSen-AD 27

E.3 Real3D-AD 27

E.4 Anomaly-ShapeNet 28

A Additional Implementation Details

A.1 Geometric Feature and Reference Construction

Geometric feature extraction. We extract region-level geometric representations using multi-scale Fast Point Feature Histogram (FPFH) descriptors. Surface normals are first estimated for each input point using at most 10 neighboring points. We then compute FPFH descriptors with three neighborhood configurations controlled by maximum neighbor counts h, $2h$, and $3h$. The search radii are set sufficiently large so that the effective neighborhood sizes are determined by these maximum-neighbor limits.

As shown in Table 4, for each point cloud, we sample $n = 4096$ region centers using farthest point sampling. For every center, we retrieve its $g = 128$ spatially nearest input points and aggregate their multi-scale geometric descriptors to obtain a region-level feature $\mathbf{f}_i$ associated with center $\mathbf{c}_i$.

Normal memory bank. For each object category, region features extracted from all normal training samples are initially pooled into a full memory bank,

$$\mathcal{M} = \bigcup_{X \in \mathcal{D}_{\mathrm{tr}}} \mathcal{F}(X). \tag{12}$$

To reduce memory consumption and matching cost, we retain 5% of the pooled features using greedy coreset selection after sparse random projection. The random projection uses an approximation parameter of 0.9. The resulting compressed feature set is used as the normal memory bank $\mathcal{M} = \{\mathbf{m}_j\}_{j=1}^{K}$.

The baseline anomaly score of test region i is the Euclidean distance to its nearest memory feature,

$$s_i^{\mathrm{base}} = \min_{\mathbf{m}_j \in \mathcal{M}} \left\| \widehat{\mathbf{f}}_i - \widehat{\mathbf{m}}_j \right\|_2. \tag{13}$$

No anomalous samples or test-set statistics are used to construct the memory bank.

Local test and normal reference sets. For each query region i, we construct a local test set using the Euclidean distances between region centers:

$$\mathcal{P}_i = \left\{ \widehat{\mathbf{f}}_\ell \;\middle|\; \mathbf{c}_\ell \in \mathrm{kNN}_{k_l}\left(\mathbf{c}_i; \{\mathbf{c}_q\}_{q=1}^{n}\right) \right\}. \tag{14}$$

Because the query center has zero distance to itself, $\mathcal{P}_i$ includes $\widehat{\mathbf{f}}_i$. We use $k_l = 7$ in the final configuration.

The query-conditioned normal reference set is constructed by exact feature-space nearest-neighbor retrieval:

$$\mathcal{Q}_i = \mathrm{kNN}_{k_r}\left(\widehat{\mathbf{f}}_i; \widehat{\mathcal{M}}\right), \tag{15}$$

where Euclidean distance is used and $k_r = 8$. Thus, $\mathcal{P}_i$ represents the spatial feature context of the query, whereas $\mathcal{Q}_i$ provides its closest normal feature anchors. Both sets are constructed independently for each query region and are subsequently passed to AKDM.

The sensitivity analysis reveals dataset-specific local optima rather than a universally optimal parameter setting. In particular, $k_r = 4$ performs favorably on the two datasets illustrated in Fig. 4. However, the optimal neighborhood and reference sizes can vary with point-cloud density, sensor noise, and defect scale. We therefore adopt a fixed configuration of $k_l = 7$ and $k_r = 8$ across all datasets, avoiding dataset-specific test-time tuning while providing a more stable estimate of both the local test distribution and the query-conditioned normal reference distribution. Moreover, these values introduce only modest computational overhead, as they remain substantially smaller than the numbers of sampled regions and memory-bank features.

A.2 Adaptive Kernel Distribution Matching

Empirical MMD Estimator. For each test region i, AKDM measures the discrepancy between the local test set $\mathcal{P}_i$ and its query-conditioned normal reference set $\mathcal{Q}_i$ using maximum mean dis-

Table 4: Implementation parameters for geometric feature and reference construction. The FPFH parameter h denotes the maximum neighbor count at the smallest descriptor scale.

Dataset	n	g	h	k_l	k_r
MiniShiftAD	4096	128	40	7	8
Anomaly-ShapeNet	4096	128	40	7	8
Real3D-AD	4096	128	40	7	8
MulSen-AD	4096	128	40	7	8

crepancy (MMD). Given a kernel $k(\cdot,\cdot)$, the biased empirical estimator is defined as

$$\begin{aligned} d^2_{\mathrm{mmd}}(\mathcal{P}_i,\mathcal{Q}_i;k) = &\frac{1}{|\mathcal{P}_i|^2}\sum_{\mathbf{x},\mathbf{x}'\in\mathcal{P}_i} k(\mathbf{x},\mathbf{x}') \\ &+\frac{1}{|\mathcal{Q}_i|^2}\sum_{\mathbf{y},\mathbf{y}'\in\mathcal{Q}_i} k(\mathbf{y},\mathbf{y}') \\ &-\frac{2}{|\mathcal{P}_i||\mathcal{Q}_i|}\sum_{\mathbf{x}\in\mathcal{P}_i}\sum_{\mathbf{y}\in\mathcal{Q}_i} k(\mathbf{x},\mathbf{y}). \end{aligned} \tag{16}$$

We adopt the biased estimator because it retains the diagonal kernel terms and yields a non-negative population-consistent estimate in practice. The separate normalization factors also allow $\mathcal{P}_i$ and $\mathcal{Q}_i$ to contain different numbers of samples.

Numerically Stable Distribution Score. Using the adaptive ARD kernel k_i^{ARD} defined in the main paper, the distribution-level anomaly score is computed as

$$s_i^{\mathrm{dist}} = \sqrt{\max\left[d^2_{\mathrm{mmd}}\left(\mathcal{P}_i,\mathcal{Q}_i;k_i^{\mathrm{ARD}}\right),0\right]+\epsilon}. \tag{17}$$

Although the biased MMD estimator is theoretically non-negative, small negative values may occasionally arise from floating-point rounding errors. We therefore clamp the estimated squared discrepancy to zero before taking the square root. The constant $\epsilon > 0$ further improves numerical stability when the discrepancy approaches zero.

Dimension-wise Relevance Decomposition. For completeness, the relevance statistic used in AKDM can be equivalently decomposed into a standardized distribution shift and a local test-side variation term:

$$\delta_{i,r} = \frac{\left(\mu^P_{i,r}-\mu^Q_{i,r}\right)^2}{v^Q_{i,r}+\epsilon}, \qquad u_{i,r} = \frac{v^P_{i,r}}{v^Q_{i,r}+\epsilon}, \tag{18}$$

followed by

$$\rho_{i,r} = \frac{a+\frac{1}{2}\delta_{i,r}}{b+\omega u_{i,r}+\epsilon}. \tag{19}$$

Here, $\delta_{i,r}$ measures the deviation of the local test features from the normal reference distribution relative to normal variation, whereas $u_{i,r}$ characterizes local test-side variation. Therefore, a feature dimension receives a large relevance value when it exhibits a pronounced distribution shift while remaining locally stable.

Score Normalization and Fusion. Before fusion with the baseline anomaly score, the distribution scores are normalized within each test sample using min–max normalization:

$$\widetilde{s}_i^{\mathrm{dist}} = \frac{s_i^{\mathrm{dist}}-\min_j s_j^{\mathrm{dist}}}{\max_j s_j^{\mathrm{dist}}-\min_j s_j^{\mathrm{dist}}+\epsilon}. \tag{20}$$

The normalized distribution score is then combined with the original feature-matching score as

$$s_i^{\mathrm{AKDM}} = s_i^{\mathrm{base}} + \lambda\widetilde{s}_i^{\mathrm{dist}}. \tag{21}$$

This additive formulation preserves the anomaly evidence provided by the original nearest-neighbor matcher while introducing a complementary correction derived from local distribution discrepancies.

Table 5: Additional parameters used for all datasets.

Component	Parameter	Value
AKDM	Local test neighbors k_l	7
	Normal reference features k_r	8
	Relevance constants a, b	1.0, 1.0
	Variation weight ω	0.5
	Softmax temperature τ	0.7
	Kernel bandwidth σ	Adaptive
	Fusion weight λ	0.4
APSP	Propagation neighbors k_p	16
	Spatial bandwidth σ_c	Adaptive
	Feature bandwidth σ_f	Adaptive
	Propagation strength α	0.4
	Correction strength β	0.3
	Propagation iterations T	4
Prediction	Object-level top-K size K_o	80

A.3 Parameter Settings

The geometric and reference-set parameters are reported in Table 4. Table 5 lists the remaining settings used for all datasets. Bandwidths set to zero in the command line are estimated adaptively from the corresponding feature or neighborhood distances.

Computational environment. All experiments are conducted on an NVIDIA GeForce RTX 4090D GPU. The software environment uses Python 3.8, PyTorch 1.12.1, and CUDA Toolkit 11.3. Open3D Zhou et al. (2018) is used for geometric feature computation, together with the PointNet++ and KNN CUDA extensions required for point-cloud processing.

A.4 Overall Algorithm

Algorithm 1 Overall Framework of the Proposed 3D Anomaly Detection Method

Require: Normal training set $\mathcal{D}_{\mathrm{tr}}$; test point cloud X^{te}; feature extractor $\mathcal{E}$; neighborhood sizes k_l, k_r, k_p; temperature τ; kernel bandwidth σ; fusion weight λ; propagation parameters α, β, T
Ensure: Point-level anomaly map $\mathbf{S}^{\mathrm{pt}}$; object-level anomaly score s^{obj}

// Stage I: Normal memory-bank construction
Initialize the normal feature memory bank $\mathcal{M} \leftarrow \varnothing$
for each normal point cloud $X \in \mathcal{D}_{\mathrm{tr}}$ **do**
  $(\mathcal{F}, \mathcal{C}) \leftarrow \mathcal{E}(X)$ ▷ MSND feature extraction and LFSA aggregation
  $\mathcal{M} \leftarrow \mathcal{M} \cup \mathcal{F}$
end for
$\mathcal{M} \leftarrow \mathrm{CoresetSelection}(\mathcal{M})$
Compute feature-wise mean and standard deviation $(\mu_{\mathcal{M}}, \sigma_{\mathcal{M}})$ from $\mathcal{M}$
Standardize the memory features: $\widehat{\mathcal{M}} \leftarrow (\mathcal{M} - \boldsymbol{\mu}_{\boldsymbol{\mathcal{M}}})/(\boldsymbol{\sigma}_{\boldsymbol{\mathcal{M}}} + \epsilon)$

// Stage II: Test feature extraction and baseline scoring
$(\mathcal{F}^{\mathrm{te}}, \mathcal{C}) \leftarrow \mathcal{E}(X^{\mathrm{te}})$
Standardize test features: $\widehat{\mathcal{F}}^{\mathrm{te}} \leftarrow (\mathcal{F}^{\mathrm{te}} - \boldsymbol{\mu}_{\boldsymbol{\mathcal{M}}})/(\boldsymbol{\sigma}_{\boldsymbol{\mathcal{M}}} + \epsilon)$
for each test region $i = 1, \ldots, n$ **do**
  $s_i^{\mathrm{base}} \leftarrow \min_{\widehat{\mathbf{m}} \in \widehat{\mathcal{M}}} \|\widehat{\mathbf{f}}_i - \widehat{\mathbf{m}}\|_2$
end for

// Stage III: Local Reference Construction and AKDM
for each test region $i = 1, \ldots, n$ **do**
  $\mathcal{P}_i \leftarrow \{\hat{f}_\ell \mid c_\ell \in \mathrm{SpatialKNN}(c_i, \mathcal{C}, k_l)\}$ ▷ Local test feature set
  $\mathcal{Q}_i \leftarrow \mathrm{FeatureKNN}(\widehat{\mathbf{f}}_i, \widehat{\mathcal{M}}, k_r)$ ▷ Query-conditioned normal reference set
  Compute $\rho_{i,r}$ from the means and variances of $\mathcal{P}_i$ and $\mathcal{Q}_i$ using Eqs. (2)–(3)
  $\mathbf{w}_i \leftarrow \mathrm{Softmax}(\boldsymbol{\rho}_i/\tau)$
  Define the query-adaptive ARD kernel:
  $k_i(\mathbf{x}, \mathbf{y}) \leftarrow \exp\Big(-\frac{1}{2\sigma^2} \sum_{r=1}^{d} w_{i,r}(x_r - y_r)^2\Big)$
  $d_i \leftarrow \sqrt{\max\big[\mathrm{MMD}^2(\mathcal{P}_i, \mathcal{Q}_i; k_i), 0\big] + \epsilon}$
end for
$\widetilde{\mathbf{d}} \leftarrow \mathrm{MinMaxNormalize}(\mathbf{d})$
$s_i^{\mathrm{AKDM}} \leftarrow s_i^{\mathrm{base}} + \lambda \widetilde{d}_i$, for all i

// Stage IV: Anomaly Probability Surface Propagation
$\mathbf{p}^{(0)} \leftarrow \mathrm{MinMaxNormalize}(\mathbf{s}^{\mathrm{AKDM}})$
for each region $i = 1, \ldots, n$ **do**
  $\mathcal{N}_i \leftarrow \mathrm{SpatialKNN}(c_i, \mathcal{C} \setminus \{c_i\}, k_p)$
  for each $j \in \mathcal{N}_i$ **do**
    $D_i(j) \leftarrow \dfrac{\|\mathbf{c}_i - \mathbf{c}_j\|_2^2}{\sigma_c^2} + \dfrac{\sum_{r=1}^{d} w_{i,r}(\widehat{f}_{i,r} - \widehat{f}_{j,r})^2}{\sigma_f^2}$
    $\kappa_{ij} \leftarrow \exp[-D_i(j)/2]$
  end for
  $a_{ij} \leftarrow \kappa_{ij} / \sum_{\ell \in \mathcal{N}_i} \kappa_{i\ell}$
end for
for $t = 0, \ldots, T-1$ **do**
  for each region $i = 1, \ldots, n$ **do**
    $p_i^{(t+1)} \leftarrow (1-\alpha) p_i^{(0)} + \alpha \sum_{j \in \mathcal{N}_i} a_{ij} p_j^{(t)}$
  end for
end for
$s_i^{\mathrm{APSP}} \leftarrow s_i^{\mathrm{AKDM}} \left[1 + \beta \left(p_i^{(T)} - p_i^{(0)}\right)\right]$

// Stage V: Point-level prediction
$\mathbf{S}^{\mathrm{pt}} \leftarrow \mathrm{Interpolate}(\{\mathbf{c}_i, s_i^{\mathrm{APSP}}\}_{i=1}^{n}, X^{\mathrm{te}})$
$s^{\mathrm{obj}} \leftarrow \mathrm{TopKMean}(S^{\mathrm{pt}}, K_o)$
return $\mathbf{S}^{\mathrm{pt}}, s^{\mathrm{obj}}$

B Theoretical Properties of APSP

This section analyzes the boundedness and convergence of the proposed Anomaly Probability Surface Propagation (APSP). We also establish the numerical range of its residual score correction.

B.1 Matrix Form of Probability-Surface Evolution

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ denote the normalized conductivity matrix, whose entries are defined as

$$A_{ij} = \begin{cases} a_{ij}, & j \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

Because $\kappa_{ij} > 0$ and the conductivities are normalized within each neighborhood, we have

$$A_{ij} \geq 0, \qquad \sum_{j=1}^{n} A_{ij} = 1. \tag{23}$$

Thus, each row of $\mathbf{A}$ defines a normalized combination of the neighboring anomaly likelihoods. The conductivity matrix need not be symmetric because its coefficients depend on the query-specific ARD weights.

Using the flow definition in Eq. 8, the accumulated flow at region i can be written as

$$\begin{aligned} \sum_{j \in \mathcal{N}_i} \Phi_{j \to i}^{(t)} &= \sum_{j \in \mathcal{N}_i} a_{ij} \left(p_j^{(t)} - p_i^{(t)} \right) \\ &= \sum_{j \in \mathcal{N}_i} a_{ij} p_j^{(t)} - p_i^{(t)}, \end{aligned} \tag{24}$$

where the second equality follows from $\sum_{j \in \mathcal{N}_i} a_{ij} = 1$. Substituting Eq. 24 into Eq. 9 gives the equivalent matrix form

$$\mathbf{p}^{(t+1)} = \alpha \mathbf{A} \mathbf{p}^{(t)} + (1 - \alpha) \mathbf{p}^{(0)}. \tag{25}$$

B.2 Boundedness of the Probability Surface

Assume that the initial relative anomaly likelihoods satisfy $\mathbf{p}^{(0)} \in [0, 1]^n$.

Proposition 1. For any $0 \leq \alpha < 1$, the APSP evolution satisfies $\mathbf{p}^{(t)} \in [0, 1]^n$ for every $t \geq 0$.

Proof. Suppose that $\mathbf{p}^{(t)} \in [0, 1]^n$. Since $\mathbf{A}$ has non-negative entries and unit row sums, each element of $\mathbf{A}\mathbf{p}^{(t)}$ is a convex combination of neighboring likelihoods. Therefore,

$$\mathbf{A}\mathbf{p}^{(t)} \in [0, 1]^n. \tag{26}$$

Eq. 25 combines $\mathbf{A}\mathbf{p}^{(t)}$ and $\mathbf{p}^{(0)}$ using the non-negative coefficients α and $1 - \alpha$, which sum to one. Hence, $\mathbf{p}^{(t+1)} \in [0, 1]^n$. The result follows by induction from $\mathbf{p}^{(0)} \in [0, 1]^n$. □

This property prevents the probability-surface evolution from producing unbounded anomaly likelihoods.

B.3 Convergence and Unique Equilibrium

Define the evolution operator

$$\mathcal{T}(\mathbf{p}) = \alpha \mathbf{A}\mathbf{p} + (1 - \alpha) \mathbf{p}^{(0)}. \tag{27}$$

Proposition 2. For $0 \leq \alpha < 1$, $\mathcal{T}$ is a contraction under the infinity norm. Consequently, APSP converges to a unique equilibrium.

Proof. For any two likelihood surfaces $\mathbf{p}$ and $\mathbf{q}$,

$$\begin{aligned} \|\mathcal{T}(\mathbf{p}) - \mathcal{T}(\mathbf{q})\|_\infty &= \alpha \|\mathbf{A}(\mathbf{p} - \mathbf{q})\|_\infty \\ &\leq \alpha \|\mathbf{p} - \mathbf{q}\|_\infty , \end{aligned} \tag{28}$$

where the inequality follows from $\|\mathbf{A}\|_\infty = 1$. Since $\alpha < 1$, the operator is contractive. The contraction mapping theorem therefore guarantees a unique equilibrium $\mathbf{p}^*$ and convergence from any initialization. □

The equilibrium satisfies

$$\mathbf{p}^* = \alpha \mathbf{A}\mathbf{p}^* + (1-\alpha)\mathbf{p}^{(0)}. \tag{29}$$

Because the spectral radius satisfies $\rho(\alpha\mathbf{A}) \leq \alpha < 1$, $\mathbf{I} - \alpha\mathbf{A}$ is invertible. The equilibrium can therefore be expressed as

$$\mathbf{p}^* = (1-\alpha)(\mathbf{I} - \alpha\mathbf{A})^{-1}\mathbf{p}^{(0)}. \tag{30}$$

Equivalently,

$$\mathbf{p}^* = (1-\alpha)\sum_{q=0}^{\infty} \alpha^q \mathbf{A}^q \mathbf{p}^{(0)}. \tag{31}$$

Eq. 31 shows that the equilibrium integrates structural support over progressively larger neighborhoods, while the influence of longer propagation paths decays geometrically.

The convergence rate is bounded by

$$\left\|\mathbf{p}^{(t)} - \mathbf{p}^*\right\|_\infty \leq \alpha^t \left\|\mathbf{p}^{(0)} - \mathbf{p}^*\right\|_\infty. \tag{32}$$

Thus, a smaller α retains more initial AKDM evidence and converges faster, whereas a larger α permits broader structural propagation.

B.4 Bounded Residual Score Correction

The final APSP score is defined as

$$s_i^{\mathrm{APSP}} = s_i^{\mathrm{AKDM}} \left[1 + \beta\left(p_i^{(T)} - p_i^{(0)}\right)\right], \qquad 0 \leq \beta \leq 1. \tag{33}$$

Proposition 3. If $s_i^{\mathrm{AKDM}} \geq 0$, then $s_i^{\mathrm{APSP}} \geq 0$, and the relative correction is bounded by β.

Proof. Proposition 1 gives $p_i^{(T)}, p_i^{(0)} \in [0,1]$, and hence

$$-1 \leq p_i^{(T)} - p_i^{(0)} \leq 1. \tag{34}$$

Therefore, the multiplicative modulation factor satisfies

$$1 - \beta \leq 1 + \beta\left(p_i^{(T)} - p_i^{(0)}\right) \leq 1 + \beta. \tag{35}$$

Because $0 \leq \beta \leq 1$, this factor is non-negative. Moreover,

$$\left|s_i^{\mathrm{APSP}} - s_i^{\mathrm{AKDM}}\right| \leq \beta s_i^{\mathrm{AKDM}}. \tag{36}$$

Thus, APSP preserves score non-negativity and prevents the structural correction from overwhelming the original AKDM evidence.

B.5 Interpretation

APSP evolves a surface of normalized relative anomaly likelihoods. Its query-adaptive conductivities need not be reciprocal, and the total likelihood is not required to remain constant. Instead, the bounded evolution redistributes local anomaly evidence according to spatial proximity and anomaly-relevant feature consistency while retaining fidelity to the initial AKDM response.

Under the assumptions that A is a nonnegative row-stochastic matrix and $0 \leq \alpha < 1$, the APSP update preserves the unit hypercube $[0,1]^n$. Moreover, the corresponding update operator is an α-contraction under the ℓ_∞ norm and therefore converges to the unique equilibrium

$$p^\star = (1-\alpha)(I - \alpha A)^{-1}p^{(0)}. \tag{37}$$

Using the Neumann-series expansion,

$$(1-\alpha)(I-\alpha A)^{-1} = (1-\alpha)\sum_{q=0}^{\infty} \alpha^q A^q, \tag{38}$$

the equilibrium can be interpreted as aggregating structural evidence along propagation paths of increasing order. In particular, the contribution of a q-hop propagation path is weighted by α^q, and therefore decays geometrically with path length.

Separately, the finite-step approximation satisfies

$$\|p^{(t)} - p^\star\|_\infty \le \alpha^t \|p^{(0)} - p^\star\|_\infty \le \alpha^t, \tag{39}$$

showing that the convergence error also decreases geometrically with the number of propagation iterations.

The equilibrium mapping is furthermore non-expansive with respect to the initial anomaly surface. Specifically, for two initial surfaces $p^{(0)}$ and $\tilde{p}^{(0)}$,

$$\begin{aligned} \|p^\star - \tilde{p}^\star\|_\infty &\le (1-\alpha) \sum_{q=0}^{\infty} \alpha^q \|A^q\|_\infty \|p^{(0)} - \tilde{p}^{(0)}\|_\infty \\ &\le \|p^{(0)} - \tilde{p}^{(0)}\|_\infty, \end{aligned} \tag{40}$$

where $\|A^q\|_\infty = 1$ follows from the row-stochasticity of A. Thus, perturbations in the initial anomaly scores are not amplified by the equilibrium propagation.

Finally, the residual correction remains bounded as

$$(1-\beta) s_i^{\mathrm{AKDM}} \le s_i^{\mathrm{APSP}} \le (1+\beta) s_i^{\mathrm{AKDM}}, \tag{41}$$

which preserves score non-negativity and limits the relative modification introduced by APSP.

C Extended Category-wise Results

C.1 Anomaly-ShapeNet

Table 6 compares the object- and point-level AUROC for all 40 categories. We retain reported aggregate results when category-wise values were not made public, and do not infer missing values.

Our method achieves 88.5% O-AUROC and 93.9% P-AUROC, surpassing the next-best results by 0.9 and 1.0 percentage points, respectively. Among the 40 categories, 34 obtain P-AUROC scores above 90%, demonstrating stable localization across diverse shapes and synthetic anomaly patterns. Perfect object-level detection is achieved on Ashtray0, Cup0, Headset0, and Microphone0. The relatively lower localization results on Bottle1, Bottle3, Bucket0, and Tap1 indicate that several categories with more ambiguous geometric variations remain challenging.

C.2 MulSen-AD

Table 7 reports the 3D-only object- and point-level AUROC for all 15 MulSen-AD categories. Category-wise results for the comparison methods follow the evaluation reported by Simple3D Cheng et al. (2026b).

Our method achieves 90.5% O-AUROC and 83.2% P-AUROC, surpassing the next-best results by 2.3 and 2.9 percentage points, respectively. It attains perfect object-level detection on Spring pad and Flat pad, while seven categories achieve P-AUROC scores of at least 93.8%. Strong detection and localization performance is also observed on Cube, Piggy, Nut, Plastic cylinder, and Light. In contrast, the results on Screen and Zipper indicate that precise defect localization remains more challenging than object-level anomaly detection for certain categories.

C.3 Real3D-AD

Table 8 reports object- and point-level AUROC for all 12 Real3D-AD categories. We retain the metrics made public for each method and do not infer unreported point-level results.

Our method achieves 81.2% O-AUROC and 93.0% P-AUROC, attaining the best overall point-level performance while remaining competitive in object-level detection. P-AUROC exceeds 90% in nine of the twelve categories, with particularly strong results on Candybar, Diamond, Duck, Fish,

Table 6: **Category-wise comparison on Anomaly-ShapeNet.** Each cell reports O-AUROC/P-AUROC (%). Bold and underlined values are the best and second-best available results, respectively. A dash denotes a category-wise metric that was not publicly reported. The Mean column retains the aggregate values reported by each paper.

Method	Ashtray0	Bag0	Bottle0	Bottle1	Bottle3	Bowl0	Bowl1	Bowl2	Bowl3	Bowl4	Mean
PatchCore	58.7/59.7	57.1/57.4	60.4/65.4	66.7/68.7	57.2/51.2	50.4/52.4	63.9/53.1	61.5/62.5	53.7/32.7	49.4/72.0	56.8/58.0
Reg3D-AD	59.7/69.8	70.6/71.5	48.6/88.6	69.5/69.6	52.5/52.5	67.1/77.5	52.5/61.5	49.0/59.3	34.8/65.4	66.3/80.0	57.2/66.8
IMRNet	67.1/67.1	66.0/66.8	55.2/55.6	70.0/70.2	64.0/64.1	68.1/78.1	70.2/70.5	68.5/68.4	59.9/59.9	67.6/57.6	66.1/65.0
ISMP	–/60.3	–/74.7	–/77.0	–/56.8	–/77.5	–/85.1	–/54.6	–/73.6	–/77.3	–/74.0	–/69.1
PO3AD	**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	83.9/89.8
MC3D-AD	96.2/80.7	80.5/85.7	79.5/90.2	70.9/**86.7**	75.6/**90.2**	93.0/77.5	**97.8**/56.2	71.9/59.7	88.5/77.9	91.1/67.0	84.2/74.8
HPPF	**100.0**/–	85.2/–	87.9/–	93.8/–	91.3/–	94.8/–	88.4/–	86.7/–	86.9/–	97.1/–	87.6/–
Simple3D	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	86.0/92.9
Ours	**100.0**/95.6	81.9/94.6	78.1/94.5	**98.2**/73.2	86.7/80.7	99.6/**99.3**	92.2/**97.1**	**96.7/98.2**	**97.4/99.5**	95.2/**99.2**	**88.5/93.9**

Method	Bowl5	Bucket0	Bucket1	Cap0	Cap3	Cap4	Cap5	Cup0	Cup1	Eraser0	Mean
PatchCore	55.8/35.8	46.9/45.9	55.1/57.1	58.0/47.2	45.3/65.3	75.7/59.5	79.0/79.5	60.0/65.5	58.6/59.6	65.7/81.0	56.8/58.0
Reg3D-AD	59.3/69.1	61.0/61.9	75.2/75.2	69.3/63.2	72.5/71.8	64.3/81.5	46.7/46.7	51.0/68.5	53.8/69.8	34.3/75.5	57.2/66.8
IMRNet	71.0/71.5	58.0/58.5	77.1/77.4	73.7/71.5	77.5/70.6	65.2/75.3	65.2/74.2	64.3/64.3	75.7/68.8	54.8/54.8	66.1/65.0
ISMP	–/53.4	–/52.4	–/67.2	–/86.5	–/73.4	–/75.3	–/67.8	–/86.9	–/60.0	–/70.6	–/69.1
PO3AD	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**	83.9/89.8
MC3D-AD	75.4/56.2	89.8/**90.2**	78.4/86.8	79.3/85.4	70.1/90.3	83.5/85.8	76.1/88.2	74.3/76.3	95.2/69.4	77.6/82.0	84.2/74.8
HPPF	89.3/–	87.0/–	**83.9**/–	**90.5**/–	86.4/–	79.8/–	80.1/–	89.2/–	85.6/–	94.0/–	87.6/–
Simple3D	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	86.0/92.9
Ours	**94.7**/97.9	85.4/70.6	83.8/**96.0**	89.3/98.1	81.4/**96.7**	**93.0**/95.4	77.2/95.4	**100.0/98.1**	**99.0/95.2**	99.5/96.7	**88.5/93.9**

Method	Headset0	Headset1	Helmet0	Helmet1	Helmet2	Helmet3	Jar0	Microphone0	Shell0	Tap0	Mean
PatchCore	58.3/58.3	63.7/46.4	54.6/54.8	48.4/48.9	42.5/45.5	40.4/73.7	47.2/47.8	38.8/48.8	49.4/61.3	75.3/73.3	56.8/58.0
Reg3D-AD	53.7/58.0	61.0/62.6	60.0/60.0	38.1/62.4	61.4/82.5	36.7/62.0	59.2/59.9	41.4/59.9	68.8/68.8	67.6/58.9	57.2/66.8
IMRNet	72.0/70.5	67.6/47.6	59.7/59.8	60.0/60.4	64.1/64.4	57.3/66.3	78.0/76.5	75.5/74.2	60.3/60.5	67.6/68.1	66.1/65.0
ISMP	–/58.0	–/70.2	–/68.3	–/62.2	–/84.4	–/72.2	–/82.3	–/66.1	–/68.7	–/52.2	–/69.1
PO3AD	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	83.9/89.8
MC3D-AD	86.2/66.6	88.6/59.2	67.2/74.9	**100.0**/59.1	60.9/81.8	**97.9**/58.5	**97.1**/84.7	91.9/89.1	**84.1**/62.5	**94.5**/50.2	84.2/74.8
HPPF	81.7/–	92.8/–	78.1/–	93.7/–	**88.4**/–	83.9/–	88.7/–	80.3/–	74.1/–	88.2/–	87.6/–
Simple3D	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	86.0/92.9
Ours	**100.0/96.4**	64.8/91.7	**82.9/93.1**	82.4/**95.1**	85.2/**95.2**	70.6/**96.0**	88.1/97.5	**100.0/96.9**	81.2/**92.1**	80.3/**88.8**	**88.5/93.9**

Method	Tap1	Vase0	Vase1	Vase2	Vase3	Vase4	Vase5	Vase7	Vase8	Vase9	Mean
PatchCore	76.6/76.8	45.5/65.5	42.3/45.3	72.1/72.1	44.9/43.0	50.6/50.5	41.7/44.7	69.3/69.3	66.2/57.5	66.0/66.3	56.8/58.0
Reg3D-AD	64.1/74.1	53.3/54.8	70.2/60.2	60.5/40.5	65.0/51.1	50.0/75.5	52.0/62.4	46.2/88.1	62.0/81.1	59.4/69.4	57.2/66.8
IMRNet	69.6/69.9	53.3/53.5	75.7/68.5	61.4/61.4	70.0/40.1	52.4/52.4	67.6/68.2	63.5/59.3	63.0/63.5	59.4/69.1	66.1/65.0
ISMP	–/55.2	–/66.1	–/84.3	–/73.3	–/76.2	–/54.5	–/47.2	–/70.1	–/85.1	–/61.5	–/69.1
PO3AD	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**	83.9/89.8
MC3D-AD	**97.0**/58.4	82.1/89.7	85.7/60.8	92.9/78.1	76.1/80.0	**87.6**/77.2	97.6/58.8	93.8/57.6	67.0/87.4	73.6/76.2	84.2/74.8
HPPF	92.4/–	89.5/–	82.9/–	**96.3**/–	**84.7**/–	80.4/–	89.0/–	93.1/–	82.5/–	**87.2**/–	87.6/–
Simple3D	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
Ours	70.4/**82.2**	**99.6**/92.7	**91.4**/86.5	94.8/98.4	77.6/**94.3**	84.5/**98.9**	**99.0/98.8**	**97.6**/98.8	**85.5/98.8**	73.6/90.3	**88.5/93.9**

Gemstone, Seahorse, and Toffees. Car and Gemstone exhibit high localization accuracy despite comparatively lower object-level scores, showing that object detection and point localization do not necessarily vary together. Shell remains the most challenging category, with lower performance on both evaluation levels.

D Additional Ablation Studies

D.1 Matching Strategy

To isolate the contribution of adaptive distribution matching, we keep the feature extractor and scoring pipeline fixed and disable APSP. As shown in Table 9, direct 1-NN matching achieves 88.2%

Table 7: **Category-wise comparison on MulSen-AD.** Each cell reports O-AUROC/P-AUROC (%). Bold and underlined values denote the best and second-best results, respectively.

Method	Venue	Capsule	Cotton	Cube	Spring pad	Screw	Screen	Piggy	Nut	Mean
M3DM-PM	CVPR'23	73.1/77.7	56.8/66.3	46.3/61.3	69.8/56.8	66.3/45.3	90.6/52.9	16.4/61.7	78.3/63.1	62.8/58.7
M3DM-PB	CVPR'23	67.1/75.3	80.5/**69.9**	45.8/71.0	51.7/65.2	95.5/44.3	92.8/56.7	44.7/57.2	75.1/68.7	70.5/61.1
PatchCore-FP	CVPR'22	89.8/91.7	25.3/55.4	72.3/57.5	98.6/62.9	97.9/57.8	91.6/60.9	**100.0**/84.8	97.1/90.3	86.0/64.0
PatchCore-FP-R	CVPR'22	90.5/91.9	26.3/54.6	66.8/43.7	**100.0**/60.1	93.1/61.0	**95.0**/58.7	99.7/62.4	**98.9**/89.6	83.3/62.0
PatchCore-PM	CVPR'22	90.3/92.1	19.7/52.8	72.2/41.7	96.5/62.1	**99.7**/59.7	89.7/53.2	98.2/60.3	**98.9**/89.7	84.0/60.5
IMRNet	CVPR'24	60.1/42.3	58.5/50.7	43.2/56.6	65.1/40.1	74.2/45.6	37.8/35.2	72.9/51.2	81.2/36.9	60.1/46.7
Reg3D-AD	NeurIPS'23	91.2/87.7	43.0/52.1	56.9/62.6	95.1/80.2	97.2/54.0	64.1/46.6	86.6/63.5	79.7/80.7	74.9/64.1
GLFM	TASE'25	96.7/93.0	**81.2**/67.9	75.6/68.0	**100.0**/70.5	63.6/60.9	86.6/50.8	73.2/77.8	94.0/95.7	78.5/66.5
Simple3D	AAAI'26	**98.1/97.3**	78.8/60.2	82.7/73.2	**100.0/81.1**	86.4/**74.7**	73.8/52.9	89.6/91.6	96.9/**97.2**	88.2/80.3
AMProp3D	**Ours**	96.2/95.4	73.0/69.7	**92.9/93.8**	**100.0**/74.8	82.9/74.0	91.6/**63.2**	97.7/**94.8**	97.7/96.4	**90.5/83.2**

Method	Venue	Flat pad	Plastic cyl.	Zipper	Button cell	Toothbrush	Solar panel	Light	Mean
M3DM-PM	CVPR'23	88.5/62.6	46.2/51.0	70.1/49.6	54.9/79.7	80.3/50.1	38.5/53.9	57.9/48.0	62.8/58.7
M3DM-PB	CVPR'23	77.2/58.3	70.6/65.2	69.8/56.3	65.9/79.9	90.1/38.6	39.5/60.1	65.3/49.5	70.5/61.1
PatchCore-FP	CVPR'22	**100.0**/70.7	94.1/83.0	79.7/55.2	**91.5**/38.2	**90.5**/60.5	62.4/20.2	97.5/70.7	86.0/64.0
PatchCore-FP-R	CVPR'22	89.3/67.8	90.8/76.6	81.3/54.5	68.7/51.2	88.8/60.4	60.5/26.5	**100.0**/70.6	83.3/62.0
PatchCore-PM	CVPR'22	94.4/63.0	93.6/76.9	73.9/50.2	79.7/47.8	89.1/60.6	61.2/27.4	99.2/69.6	84.0/60.5
IMRNet	CVPR'24	71.4/54.2	62.1/41.2	63.0/49.6	70.2/48.5	61.5/51.9	34.4/53.3	45.7/41.5	60.1/46.7
Reg3D-AD	NeurIPS'23	90.8/69.2	76.5/67.0	47.0/53.6	78.2/70.6	81.2/47.2	**66.0**/60.9	89.7/65.1	74.9/64.1
GLFM	TASE'25	94.6/77.2	81.5/67.8	81.3/57.4	56.7/43.3	84.9/57.8	40.9/57.5	66.1/52.6	78.5/66.5
Simple3D	AAAI'26	**100.0**/82.0	**99.4**/96.4	88.5/53.0	85.1/80.1	82.9/**73.3**	64.4/**94.9**	96.8/**95.9**	88.2/80.3
AMProp3D	**Ours**	**100.0/83.0**	99.2/**96.8**	**95.6/60.2**	89.5/**85.1**	80.3/70.7	62.9/**94.9**	98.1/95.0	**90.5/83.2**

O-AUROC and 80.3% P-AUROC. Adaptive kNN improves these results by 0.3 and 1.1 percentage points, respectively, suggesting that using multiple normal references reduces the sensitivity to an individual nearest neighbor. Replacing point-to-point matching with local MMD further increases P-AUROC to 82.0%, corresponding to a 1.7-point improvement over direct 1-NN, while producing a smaller O-AUROC gain of 0.1 points. This difference indicates that local distribution comparison mainly benefits fine-grained anomaly localization. Incorporating the adaptive ARD kernel into local MMD yields the best performance of 89.1% O-AUROC and 82.5% P-AUROC, improving the local MMD variant by 0.8 and 0.5 points and the direct 1-NN baseline by 0.9 and 2.2 points. These controlled comparisons support the respective contributions of distribution-level matching and query-adaptive channel weighting.

D.2 Bayesian-inspired ARD

We next evaluate different feature-dimension weighting strategies under the same local MMD configuration, with APSP disabled. As reported in Table 10, uniform weighting obtains 87.9% O-AUROC and 80.5% P-AUROC. Mean-difference ARD increases the two metrics to 88.2% and 81.8%, demonstrating that assigning non-uniform weights according to channel-wise distribution shifts is more effective than treating all feature dimensions equally. The proposed Bayesian-inspired ARD further achieves 89.1% O-AUROC and 82.5% P-AUROC, exceeding uniform weighting by 1.2 and 2.0 points and mean-difference ARD by 0.9 and 0.7 points, respectively. Unlike mean-difference weighting, which considers only the magnitude of the mean shift, our formulation jointly accounts for the standardized distribution shift and relative local variation. Consequently, channels exhibiting large but unstable variations are downweighted, whereas channels with consistent test–reference discrepancies receive greater emphasis. The larger improvement in P-AUROC confirms the importance of variation-aware channel selection for local anomaly discrimination.

D.3 Spatial-feature Propagation

Finally, we investigate the spatial and feature components of APSP while keeping the AKDM scores and all other settings fixed. Without propagation, the model obtains 89.1% O-AUROC and 82.5%

Table 8: **Category-wise comparison on Real3D-AD.** Each cell reports O-AUROC/P-AUROC (%). Bold and underlined values denote the best and second-best available results, respectively. A dash indicates that the point-level result was not publicly reported. PatchCore denotes the FPFH+Raw variant used in the Real3D-AD benchmark.

Method	Venue	Airplane	Candybar	Car	Chicken	Diamond	Duck	Mean
PatchCore	CVPR'22	84.8/55.6	57.0/74.9	77.7/74.0	85.3/55.8	78.4/85.4	62.8/65.8	68.2/69.2
Reg3D-AD	NeurIPS'23	71.6/63.1	82.7/72.4	69.7/71.8	85.2/67.6	90.0/83.5	58.4/50.3	70.4/70.0
IMRNet	CVPR'24	76.2/–	75.5/–	71.1/–	78.0/–	90.5/–	51.7/–	72.5/–
ISMP	AAAI'25	**85.8**/75.3	85.2/90.7	73.1/83.6	71.4/79.8	94.8/92.6	71.2/87.6	76.7/83.6
PO3AD	CVPR'25	80.4/–	78.5/–	65.4/–	68.6/–	80.1/–	82.0/–	76.5/–
MC3D-AD	IJCAI'25	85.0/62.8	83.0/91.0	74.9/81.9	71.5/64.0	95.5/94.2	83.1/82.2	78.2/76.8
PASDF	ICCV'25	62.8/77.7	78.8/54.6	95.9/80.2	73.9/76.8	89.4/69.9	65.8/78.2	80.2/74.5
Reg2Inv	NeurIPS'25	81.8/**92.3**	**100.0**/96.9	75.8/94.4	**94.4/91.0**	**100.0**/97.9	75.0/93.7	78.0/87.8
Simple3D	AAAI'26	76.5/88.1	65.1/96.2	**98.1/99.2**	82.6/86.1	**100.0**/99.0	77.8/96.6	80.4/92.3
AMProp3D	**Ours**	76.5/90.4	94.9/**98.9**	62.8/96.4	83.8/86.7	99.7/**99.5**	**84.6/97.4**	**81.2/93.0**

Method	Venue	Fish	Gemstone	Seahorse	Shell	Starfish	Toffees	Mean
PatchCore	CVPR'22	83.7/78.1	35.9/53.9	76.7/80.8	66.3/75.3	47.1/61.3	62.6/54.9	68.2/69.2
Reg3D-AD	NeurIPS'23	91.5/82.6	41.7/54.5	76.2/81.7	58.3/81.1	50.6/61.7	68.5/75.9	70.4/70.0
IMRNet	CVPR'24	88.0/–	67.4/–	60.4/–	66.5/–	67.4/–	77.4/–	72.5/–
ISMP	AAAI'25	94.5/88.6	46.8/85.7	72.9/81.3	62.3/83.9	66.0/64.1	84.2/89.5	76.7/83.6
PO3AD	CVPR'25	85.9/–	69.3/–	75.6/–	80.0/–	75.8/–	77.1/–	76.5/–
MC3D-AD	IJCAI'25	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
PASDF	ICCV'25	**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
Reg2Inv	NeurIPS'25	67.2/84.6	**73.5**/90.7	53.2/64.5	69.2/**90.6**	**84.1**/84.0	62.6/73.7	78.0/87.8
Simple3D	AAAI'26	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
AMProp3D	**Ours**	91.4/**97.9**	69.0/**97.4**	89.3/**96.1**	58.4/70.2	75.7/**87.5**	88.3/**97.9**	**81.2/93.0**

Table 9: Ablation of matching strategies on MulSen-AD. APSP is disabled. The best and second-best results are shown in bold and underlined.

Matching strategy	O-AUROC	P-AUROC
Direct 1-NN	88.2	80.3
Adaptive kNN	88.5	81.4
Local MMD	88.3	82.0
AKDM	**89.1**	**82.5**

P-AUROC. Spatial-only propagation improves the results to 90.2% and 82.8%, providing gains of 1.1 and 0.3 points, respectively. By contrast, feature-only propagation achieves 89.5% O-AUROC and 83.0% P-AUROC, with a smaller object-level gain but a larger improvement in point-level localization. These different behaviors indicate that spatial proximity primarily reinforces region-level continuity, whereas feature affinity helps preserve locally consistent anomaly responses. Combining the two affinities produces the best performance of 90.5% O-AUROC and 83.2% P-AUROC, improving the non-propagation variant by 1.4 and 0.7 points. It also provides moderate but consistent gains over both single-affinity variants, supporting the complementary roles of spatial proximity and anomaly-relevant feature consistency in APSP.

D.4 Accuracy–Efficiency Analysis of AKDM and APSP

Experimental setup. We evaluate four configurations across all 15 MulSen-AD categories: the baseline, baseline with AKDM, baseline with APSP, and the full method. For each category, a class-specific memory bank is constructed from the training data and shared across all configurations.

Table 10: **Ablation of feature-dimension weighting on MulSen-AD.** All variants use the same local MMD configuration, while APSP is disabled. Δ denotes the absolute change relative to uniform weighting. Tied best results at the reported precision are shown in bold.

Weighting strategy	O-AUROC	ΔO	P-AUROC	ΔP
Uniform	87.9	0.0	80.5	0.0
Mean-difference ARD	88.2	+0.3	81.8	+1.3
Bayesian-inspired ARD	**89.1**	+1.2	**82.5**	+2.0

Table 11: **Ablation of spatial–feature propagation on MulSen-AD.** All variants use the same AKDM configuration, while only the affinity components used for propagation are changed.

Propagation strategy	O-AUROC	P-AUROC
Without propagation	89.1	82.5
Spatial only	90.2	82.8
Feature only	89.5	83.0
Spatial–feature propagation	**90.5**	**83.2**

The baseline employs one-nearest-neighbor matching without APSP. AKDM augments the baseline 1-NN score with local ARD-RBF MMD scoring using a query-conditioned normal reference set, whereas APSP propagates anomaly evidence among spatially proximal and feature-consistent regions. The full method enables both modules.

All configurations share the same feature-extraction and inference settings, including MSND, LFSA, 4,096 sampled groups, a group size of 128, top-k score interpolation, and the same score-map construction procedure. The class-specific memory banks contain an average of 18,991.6 patches.

Timing protocol and scope. We evaluate inference efficiency across all 15 categories. For each category and configuration, six timed inference calls are collected in each run, and the complete measurement is repeated three times. Each configuration is warmed up before timing. Measurements are performed in inference mode using a wall-clock timer, with CUDA synchronization immediately before and after each timed call.

Timing starts after dataloader retrieval and covers the complete `predict()` call. It therefore includes normal estimation, multi-scale FPFH extraction, farthest point sampling, local feature aggregation, memory-bank matching, the enabled AKDM and/or APSP operations, point-score interpolation, score-map construction, and transfer of predictions to host arrays. Dataloader retrieval, memory-bank construction, warm-up calls, and AUROC computation are excluded from the reported latency.

Let $T_{\mathrm{e2e},c}$ denote the mean end-to-end inference latency, in milliseconds, under configuration c. The corresponding throughput is computed as

$$\mathrm{FPS}_c = \frac{1000}{T_{\mathrm{e2e},c}}. \tag{42}$$

The relative latency overhead with respect to the baseline is defined as

$$\mathrm{Overhead}_c = 100\frac{T_{\mathrm{e2e},c} - T_{\mathrm{e2e,base}}}{T_{\mathrm{e2e,base}}}. \tag{43}$$

We additionally report the standard deviation and 95th-percentile (P95) latency to characterize runtime variability. Peak GPU memory denotes the maximum CUDA allocated memory observed during inference, expressed in MiB, rather than the total device memory usage.

Accuracy and efficiency. The accuracy results in Table 12 are obtained from the separate full accuracy ablation, whereas the efficiency statistics are measured using the repeated timing protocol described above. On MulSen-AD, the object-level anomaly score is computed as the mean of the 80 highest point-level anomaly scores. O-AUROC is calculated from the resulting object scores and object labels, while P-AUROC is computed from point-level anomaly scores and point labels; both metrics are averaged across categories.

Table 12: **Accuracy and inference efficiency on MulSen-AD.** O-AUROC and P-AUROC are obtained from the full accuracy evaluation, whereas efficiency statistics are measured using the repeated timing protocol over all 15 categories. Latency covers the complete `predict()` call after dataloader retrieval.

Configuration	O-AUROC (%)↑	P-AUROC (%)↑	Mean latency (ms)↓	Std. (ms)↓	P95 (ms)↓	FPS↑	Overhead (%)↓	Peak GPU (MiB)↓
Baseline	88.2	80.3	970.54	218.79	1358.57	1.030	0.000	1263.39
Baseline + AKDM	89.1	82.5	979.16	217.48	1356.78	1.021	0.888	1263.42
Baseline + APSP	89.4	82.1	973.10	218.60	1362.42	1.028	0.264	1263.39
Full method	**90.5**	**83.2**	982.12	218.53	1370.27	1.018	1.193	1263.42

AKDM improves O-AUROC/P-AUROC from 88.2/80.3 to 89.1/82.5, corresponding to gains of 0.9/2.2 percentage points. APSP alone achieves 89.4/82.1, yielding improvements of 1.2/1.8 points. Combining the two modules produces the highest accuracy, reaching 90.5 O-AUROC and 83.2 P-AUROC, which represents gains of 2.3 and 2.9 percentage points over the baseline, respectively.

The additional computational cost remains limited. APSP increases mean latency from 970.54 to 973.10 ms, corresponding to an overhead of only 0.264%. AKDM requires 979.16 ms, resulting in an overhead of 0.888%. The full method reaches 982.12 ms, an increase of 11.58 ms or 1.193% relative to the baseline. Accordingly, throughput changes only slightly from 1.030 FPS for the baseline to 1.018 FPS for the full method. Peak allocated GPU memory remains essentially unchanged, increasing by only approximately 0.03 MiB when AKDM is enabled. These results indicate that the improvements in anomaly detection and localization are obtained with only a modest increase in end-to-end inference cost and negligible additional GPU memory.

Category-wise runtime analysis. Table 13 further reports the runtime behavior across individual MulSen-AD categories. Baseline latency ranges from 668.66 to 1299.98 ms, corresponding to a throughput range of 1.496 to 0.769 FPS. Capsule exhibits the lowest inference latency, whereas Solar panel, Plastic cylinder, and Cube are among the most computationally demanding categories.

APSP introduces a small positive latency overhead for 13 of the 15 categories. The slight negative differences observed for Nut and Toothbrush are sufficiently small to fall within normal runtime variation rather than indicating a systematic speedup. AKDM and the full configuration are consistently slower than the baseline across all categories, but their relative overhead remains small overall. This suggests that the additional computation introduced by the two modules is relatively stable across objects with different geometric characteristics and point-cloud sizes.

Plastic cylinder exhibits noticeably larger runtime variability than the remaining categories, with a baseline standard deviation of 446.06 ms and a coefficient of variation of approximately 37.34%. Its category-level runtime should therefore be interpreted with some caution, although this variation has limited influence on the overall average trends.

E Additional Qualitative Results

This section presents additional qualitative results on MiniShiftAD, MulSen-AD, Anomaly-ShapeNet, and Real3D-AD. Each example compares the predicted point-level anomaly map (Ours) with the corresponding ground-truth annotation (GT). Warmer colors in the predicted maps indicate higher anomaly scores, while red and blue regions in GT denote anomalous and normal areas, respectively. We assess both the alignment of predicted responses with annotated defects and the suppression of activations on unaffected surfaces, including challenging cases with incomplete or diffuse localization.

Each column corresponds to an object category. Within each two-row block, the first row shows the point-level anomaly map produced by our method, while the second row shows the corresponding ground-truth annotation. In the predicted maps, warmer colors indicate higher anomaly likelihoods; in the ground truth, red and blue denote anomalous and normal regions, respectively.

Table 13: **Category-wise inference overhead on MulSen-AD.** Baseline latency and FPS are reported together with the relative end-to-end latency overhead introduced by APSP, AKDM, and their combination.

Category	Baseline latency (ms)	Baseline FPS	APSP Overhead (%)	AKDM Overhead (%)	Full Overhead (%)
Capsule	668.7	1.496	0.116	1.845	2.236
Cotton	888.2	1.126	0.103	0.333	1.056
Cube	1182.7	0.845	0.482	0.868	1.464
Spring pad	879.9	1.136	0.096	0.980	1.400
Screw	723.1	1.383	0.815	1.554	1.816
Screen	958.7	1.043	0.091	1.004	0.924
Piggy	920.7	1.086	0.446	1.247	1.185
Nut	1103.0	0.907	-0.033	1.064	0.819
Flat pad	1014.6	0.986	0.642	0.906	1.263
Plastic cylinder	1194.7	0.837	0.070	0.495	0.894
Zipper	817.5	1.223	0.695	1.346	1.518
Button cell	1174.4	0.851	0.612	0.831	1.163
Toothbrush	915.3	1.093	-0.115	0.500	0.927
Solar panel	1300.0	0.769	0.029	0.253	0.821
Light	816.6	1.225	0.014	0.883	1.114

E.1 MINISHIFTAD

Fig. 5 extends the MiniShiftAD comparison presented in the main text with twelve additional examples. For `nut` and `cube`, the strongest responses closely follow the thin annotated surface defects rather than spreading over the entire object. In `button_cell`, `capsule`, `piggy`, and `plastic_cylinder`, the anomaly maps exhibit compact response peaks around the annotated defect regions. The two `screw` examples further demonstrate different spatial patterns: one contains a small defect on the ring, whereas the other exhibits an elongated anomaly along the shaft. For `screen` and `toothbrush`, the predicted responses highlight the affected edge or tip, although weaker activations are also observed outside the annotated regions. Broader background responses remain visible for `light` and the ring-shaped `screw`, indicating that subtle surface variations may still interfere with the precise localization of small defects.

E.2 MULSEN-AD

Figure 6 presents localization results across twelve MulSen-AD categories. Compact high-score regions closely overlap the annotated defects in `button_cell`, `capsule`, `nut`, and `plastic_cylinder`. The elongated responses in `cotton` and along the shaft of `screw` also align well with the spatial extent and orientation of their defect annotations. For `spring_pad`, elevated anomaly scores are observed near both annotated ends. The examples also reveal limitations in recovering the full defect extent: the long damaged rim of `flat_pad` is captured mainly by a shorter high-response segment. More diffuse activations on normal surfaces are visible for `light` and `piggy`. Overall, compact defects tend to produce focused responses, whereas extended defects and background surface variations remain more challenging to localize precisely.

E.3 REAL3D-AD

Fig. 7 presents qualitative examples from real-scanned objects with diverse surface geometries. Localized high-score responses closely coincide with the annotated regions near the joint of `starfish`, the tail of `seahorse`, the body of `fish`, and the upper surface of `diamond`. For `shell`, both annotated defect regions are captured by elevated anomaly scores, although its ridged surface also produces diffuse lower-intensity responses. The defects in `candybar`, `toffees`, and `chicken` remain visible as localized peaks, while the latter two examples exhibit additional activations in unannotated regions. These results show that anomaly responses can remain spatially

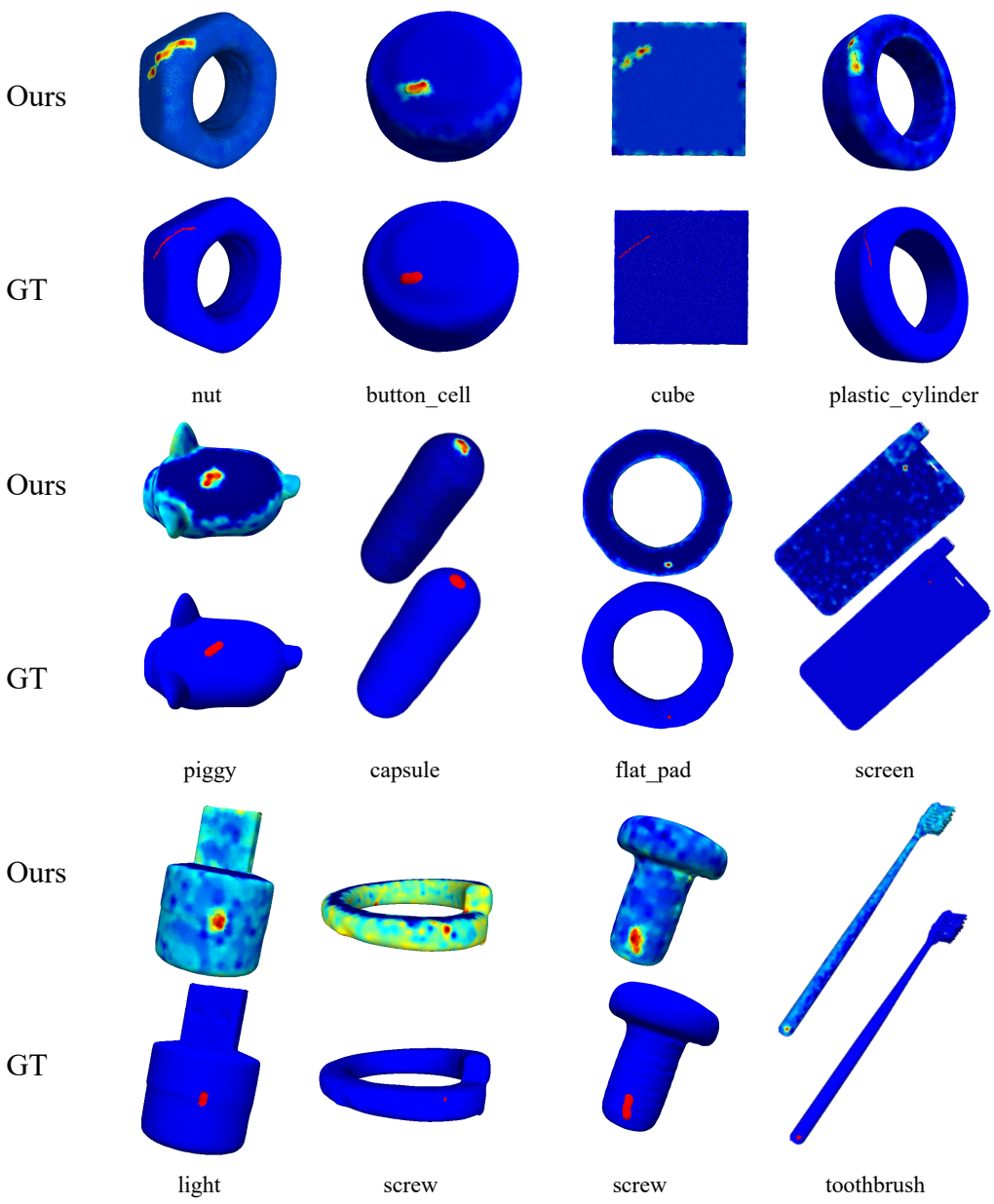

Figure 5: **Qualitative visualization on MiniShiftAD.** Warmer colors indicate higher anomaly scores.

concentrated across different scanned geometries, although surface irregularities may still broaden the predicted anomaly maps.

E.4 Anomaly-ShapeNet

Figs. 8 and 9 present qualitative results across a broader range of object shapes and defect locations, including bottles, bowls, buckets, headsets, helmets, and vases. In `bottle0`, `bottle1`, `bottle3`, and the headset examples, elevated anomaly scores are concentrated near the annotated neck or band defects. The larger defective regions in `bucket0` and `bucket1` produce broader responses, whereas the anomalies in `cup0`, `cup1`, and `shelf0` remain localized around the affected handles or corners. These examples demonstrate that the predicted responses can adapt to both small boundary defects and larger surface anomalies. However, the spatial coverage is not uni-

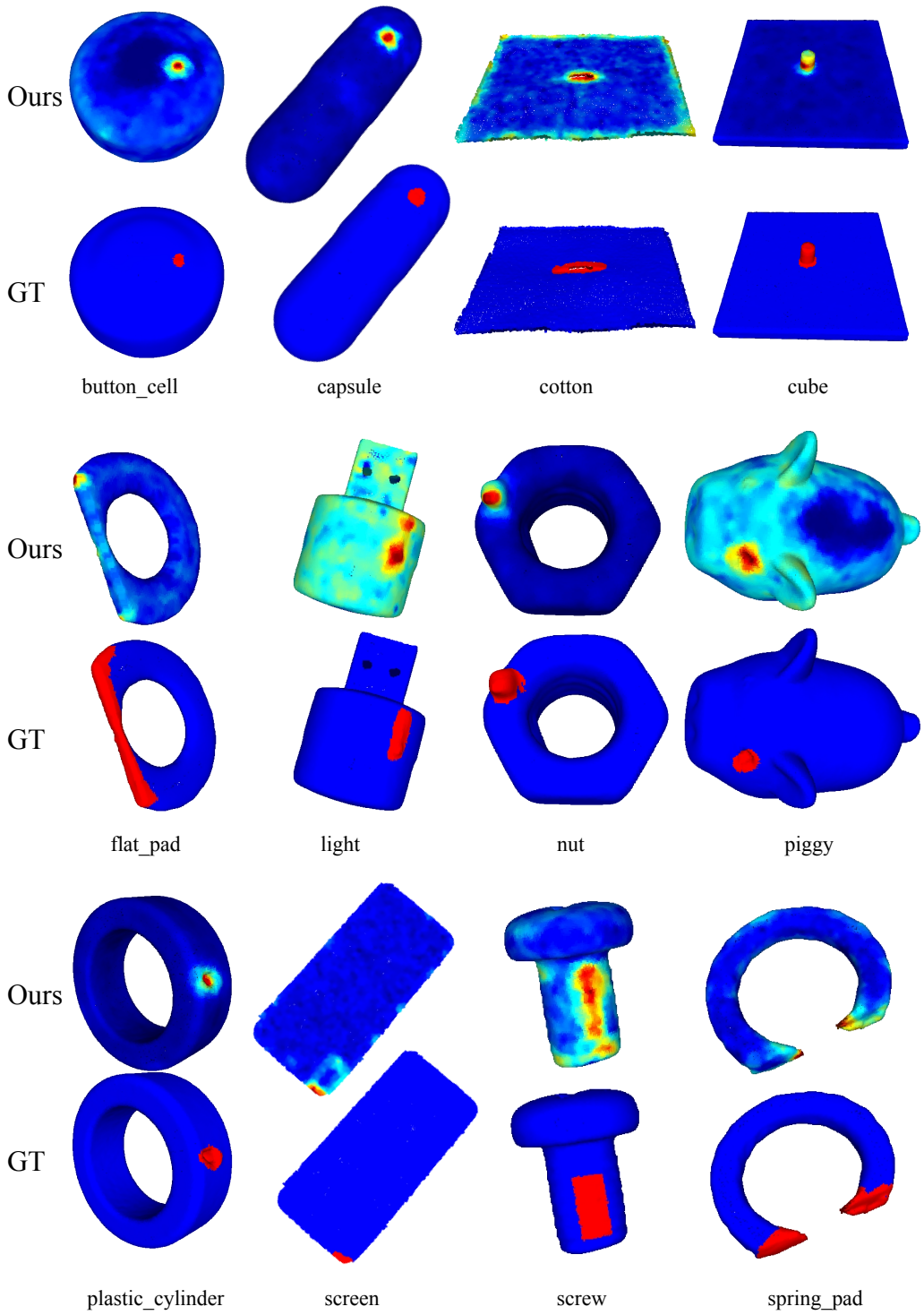

Figure 6: **Qualitative visualization on MulSen-AD.** Warmer colors indicate higher anomaly scores.

formly complete across all cases: the high-score regions in `helmet2` and `vase9` capture only part of the corresponding ground-truth defects. Low-to-moderate activations on some normal surfaces further indicate that precise defect boundaries are not always fully recovered.

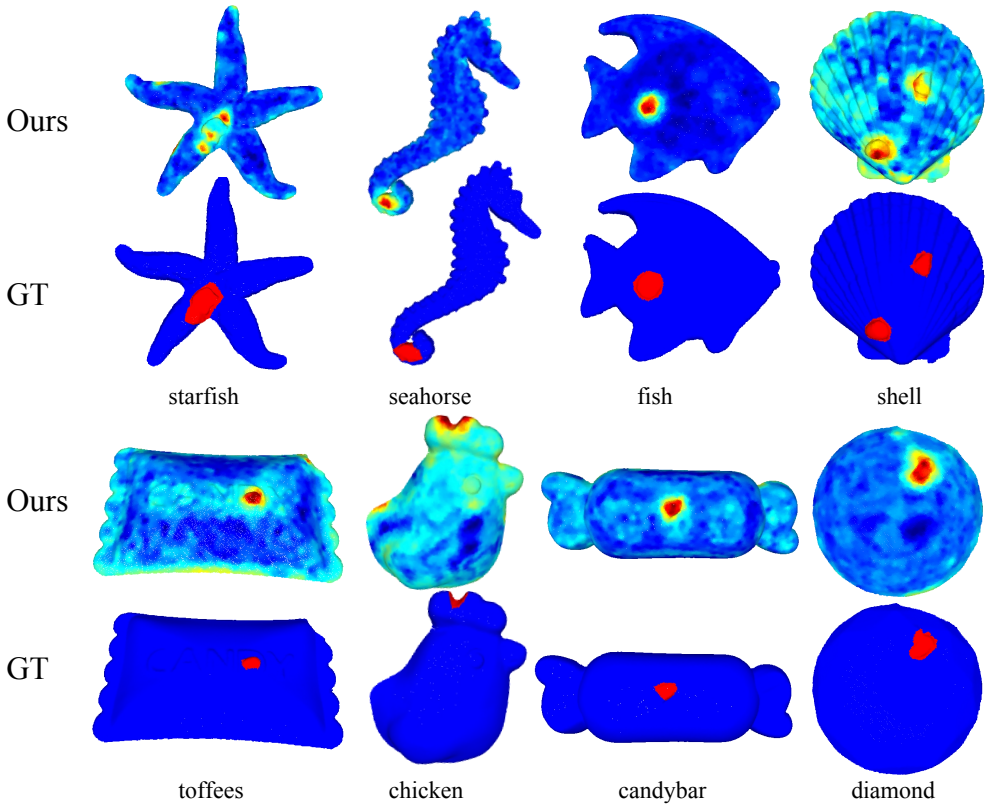

Figure 7: **Qualitative visualization on Real3D-AD.** Warmer colors indicate higher anomaly scores.

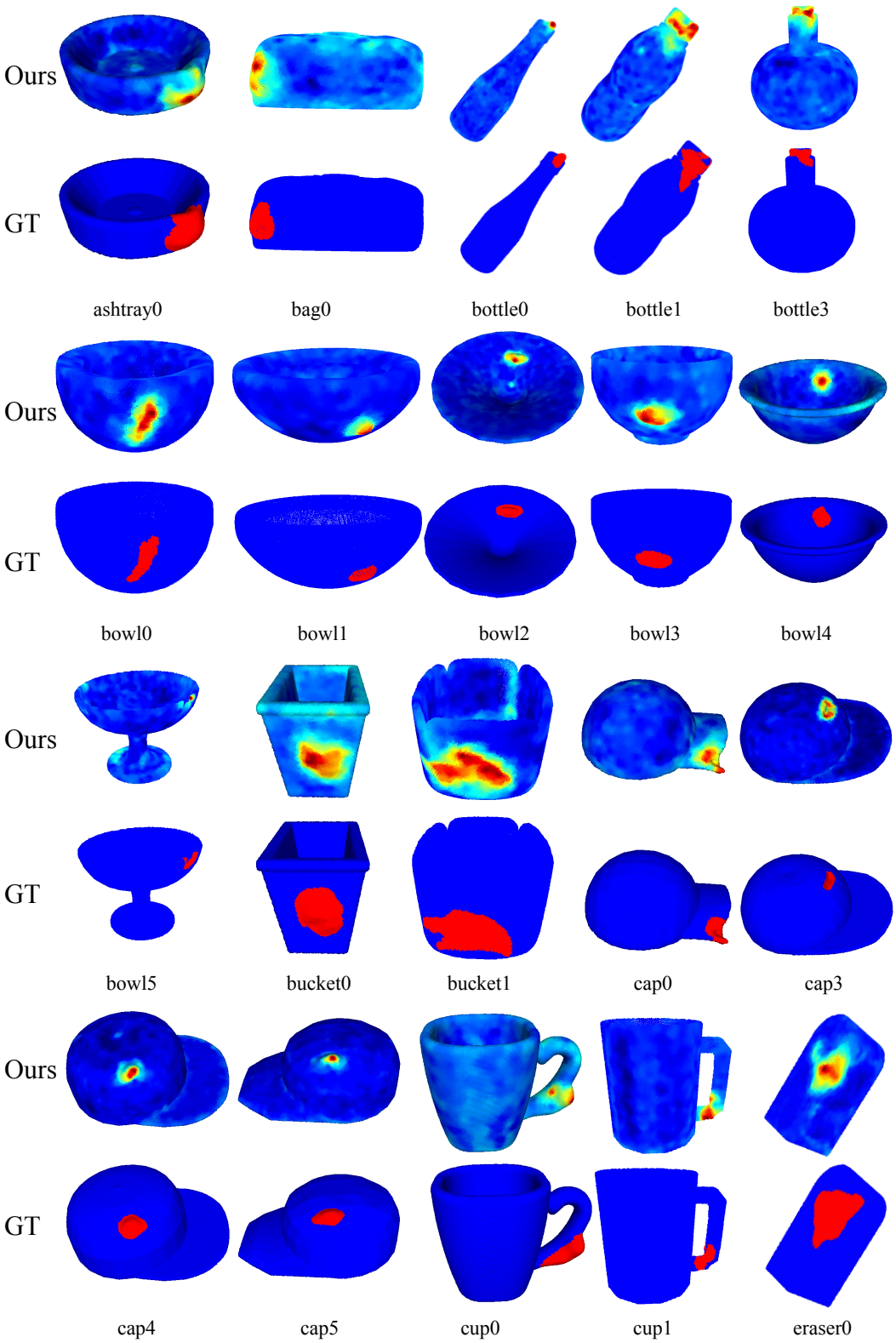

Figure 8: **Qualitative visualization on Anomaly-ShapeNet (a).** Warmer colors indicate higher anomaly scores.

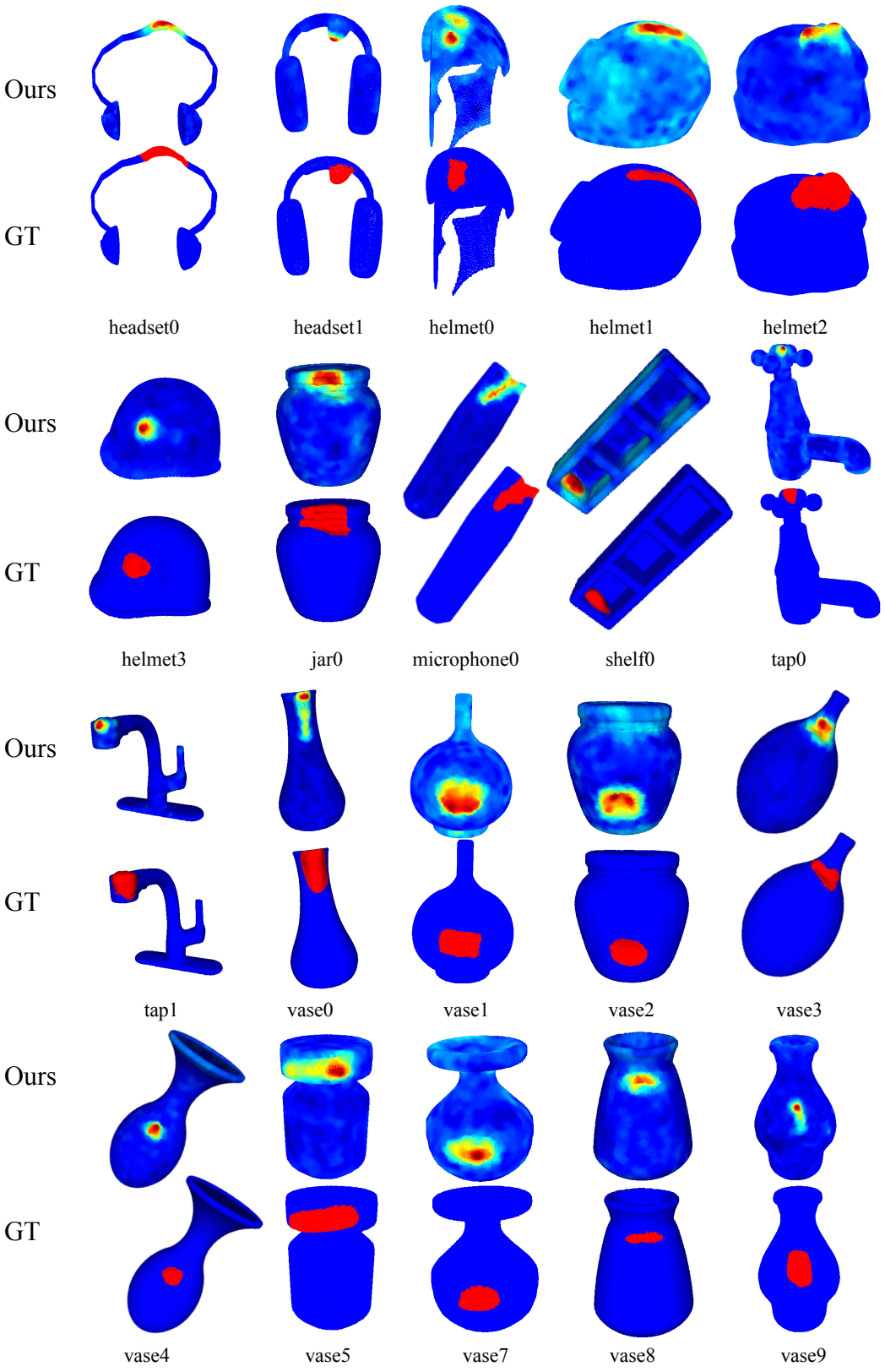

Figure 9: **Qualitative visualization on Anomaly-ShapeNet (b).** Warmer colors indicate higher anomaly scores.