Learning Anomaly-Aware Normal Subspaces for Few-Shot Industrial Anomaly Detection

Anonymous authors
Paper under double-blind review

Abstract

Few-shot industrial anomaly detection aims to identify unseen defects using only a few normal images. Recent reconstruction-based methods seek to model normality through feature restoration. However, they may fail to generalize to unseen normal variations or inadvertently reconstruct anomalous information, leading to false positives and false negatives.To address this issue, we propose NormProj, which learns an anomaly-aware normal subspace by determining which feature variations should be preserved or rejected. An orthogonal-basis generator constructs the normal subspace from augmented support features, while joint normal reconstruction and paired synthetic-anomaly restoration encourage the preservation of legitimate normal variations and the exclusion of anomaly-induced changes. After training, the generator is discarded, enabling efficient projection-based anomaly scoring without an additional neural reconstruction network. Experiments on MVTec-AD and VisA demonstrate improved image-level anomaly detection over reconstruction- and PCA-based baselines, together with competitive localization performance across multiple few-shot settings. Code is available at https://github.com/sfhiswfgh/NormProj.

1 Introduction

Few-shot industrial anomaly detection aims to identify unseen defects using only a few normal images. Recent visual foundation models provide transferable representations for this challenging setting (Oquab et al., 2024; Damm et al., 2025). Existing methods predominantly rely on feature matching, such as PatchCore (Roth et al., 2022) and AnomalyDINO (Damm et al., 2025), which identify anomalies by comparing query features with normal references. However, limited reference samples cannot fully cover legitimate normal variations, potentially causing unseen normal patterns to be misidentified as anomalies. Recently, FoundAD (Zhai et al., 2026) revisits feature reconstruction for few-shot anomaly detection by learning a nonlinear projector to restore normal features from clean or synthetically corrupted inputs. Nevertheless, accurately reconstructing observed normal patterns does not guarantee reliable reconstruction of unseen normal features or sufficient rejection of anomalous information. As illustrated in Fig. 1(a), unseen normal features may produce large reconstruction residuals, while anomalous features may be reconstructed too well, resulting in false positives and false negatives, respectively. This highlights the need to preserve unseen normal variations while maintaining sensitivity to anomalous changes. SubspaceAD (Lendering et al., 2026) approaches normality modeling through PCA-based subspace estimation. By preserving feature variations along learned normal directions, orthogonal projection can explain unseen normal variations within the subspace without predicting their specific normal counterparts. However, PCA selects directions according to normal feature variance without considering whether these directions also reconstruct anomalous changes. As illustrated in Fig. 1(b), anomaly-induced changes aligned with retained principal directions may also be preserved, leading to small projection residuals and potential false negatives. This reveals a critical limitation: the effectiveness of a normal subspace depends not only on its ability to preserve normal variations but also on which anomalous changes it excludes.

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.

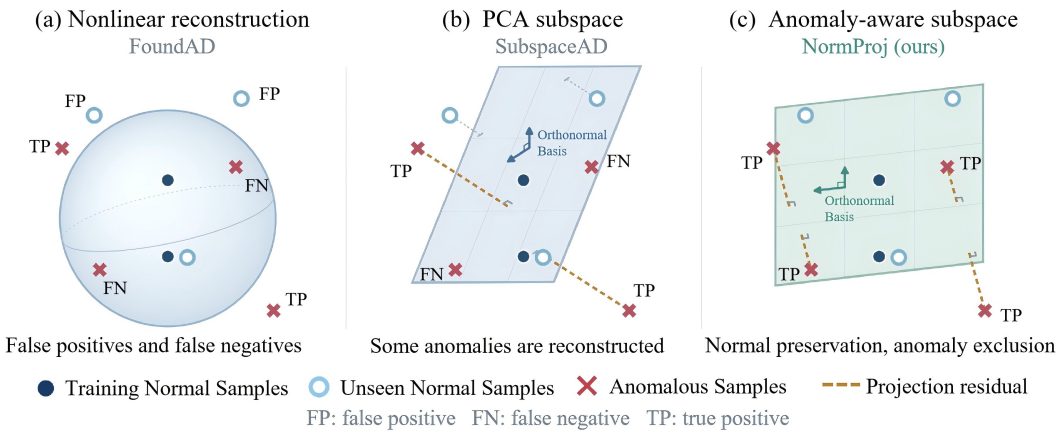

Figure 1: Schematic comparison of normality modeling. (a) Nonlinear reconstruction may poorly reconstruct unseen normal features while reconstructing anomalies too well. (b) PCA may retain anomalous changes, leading to small projection residuals. (c) NormProj learns an anomaly-aware normal subspace to preserve normal variations and exclude anomaly-induced changes.

To address this limitation, we propose NormProj, an anomaly-aware normal subspace learning framework that explicitly incorporates both normal preservation and anomaly exclusion. As illustrated in Fig. 1(c), NormProj encourages legitimate normal variations to be preserved within the learned subspace while keeping anomaly-induced changes in the projection residual. An orthogonal-basis generator constructs the subspace from augmented normal support features and is optimized through joint normal reconstruction and paired synthetic-anomaly restoration. Unlike normal-only PCA estimation, this joint objective guides the selection of normal directions using both normal and synthetic-anomaly information. Importantly, the basis generator is discarded after training. During inference, NormProj performs anomaly scoring through orthogonal projection without an additional neural reconstruction network, reducing inference-time computational overhead. Experiments on MVTec-AD (Bergmann et al., 2019) and VisA (Zou et al., 2022) demonstrate that NormProj improves image-level anomaly detection over FoundAD and SubspaceAD using the same frozen visual encoder, while maintaining competitive localization performance across multiple few-shot settings. Our main contributions are summarized as follows:

- We propose NormProj, an anomaly-aware normal subspace learning framework that preserves legitimate normal variations while discouraging anomalous changes from being reconstructed as normal.
- We develop an orthogonal-basis generator optimized through joint normal reconstruction and paired synthetic-anomaly restoration, enabling efficient projection-based inference without retaining the generator.
- Experiments on MVTec-AD and VisA demonstrate improved image-level anomaly detection over FoundAD and SubspaceAD with the same frozen encoder, alongside competitive localization performance.

2 Related Work

2.1 Few-Shot Industrial Anomaly Detection

Few-shot industrial anomaly detection aims to identify unseen defects using only a few normal samples. Feature matching methods compare query features with normal reference features. SPADE retrieves corresponding features from normal images, while PatchCore constructs a representative memory bank for nearest neighbor matching Cohen & Hoshen

(2020); Roth et al. (2022). PaDiM models normal patch feature distributions using Gaussian statistics Defard et al. (2021). AnomalyDINO exploits transferable foundation features for patch matching, while FewSOME learns a one-class similarity metric Damm et al. (2025); Belton et al. (2023). These methods directly exploit target-category information, but limited normal references may not sufficiently cover legitimate appearance variations. Methods retaining large memory banks may also incur additional storage and retrieval cost. To enrich the information available from limited normal references, other methods exploit transferable knowledge across categories or modalities. RegAD learns category-generalizable feature registration, InCTRL uses normal samples as in-context prompts, and DictAS transfers dictionary lookup capabilities to unseen categories Huang et al. (2022); Zhu & Pang (2024); Qu et al. (2025). These methods extend the available reference information, but transferred knowledge may not fully capture category-specific variations. Vision-language approaches further introduce semantic information. WinCLIP leverage text-aligned visual representations, while PromptAD and IIPAD learn prompts for anomaly discrimination Jeong et al. (2023); Li et al. (2024); Lv et al. (2025). UniVAD incorporates text-assisted component understanding and structural relationships Gu et al. (2025). Although these approaches enrich normal references with transferable or semantic knowledge, semantic priors may not adequately describe subtle local defects. Feature comparison therefore remains dependent on the coverage and suitability of reference representations.

To alleviate dependence on direct reference matching, reconstruction-based methods construct normal counterparts for query features and identify anomalies through reconstruction residuals. FastRecon employs distribution-regularized regression to reconstruct normal features from limited references, encouraging consistency with normal feature statistics. However, its reconstruction remains constrained by the estimated support distribution Fang et al. (2023). Recently, FoundAD learns a nonlinear projector through synthetic-anomaly restoration, but its reconstruction mapping may not adequately generalize to unseen normal variations or fully reject anomalous information Zhai et al. (2026). SubspaceAD estimates a PCA-based normal subspace and detects anomalies using projection residuals, avoiding a learned neural reconstruction mapping. However, normal-only PCA estimation does not explicitly penalize anomaly-induced changes aligned with the retained directions Lendering et al. (2026). These methods indicate that reliable reconstruction depends not only on the reconstruction mechanism, but also on how the normal feature space is constructed and which variations it can explain.

2.2 Normality Modeling

Learning effective representations of high-dimensional data requires determining which structures should be retained and which irrelevant information should be suppressed (Zhang et al., 2026). Classical methods address this problem through different modeling objectives. PCA selects orthogonal directions of maximum variance, while probabilistic PCA provides a probabilistic formulation and Kernel PCA captures nonlinear structures through kernel methods pca (1986); Michael et al. (1999); Schölkopf et al. (1998). Robust PCA separates low-rank structure from sparse corruption, whereas nonnegative matrix factorization learns additive components Candes et al. (2011); Lee & Seung (1999). Manifold learning methods, including Isomap, locally linear embedding, and Laplacian Eigenmaps, characterize geometric relationships among observations Tenenbaum et al. (2000); Roweis & Saul (2000); Belkin & Niyogi (2003). These methods capture different aspects of data structure, but their standard objectives do not jointly specify which variations should be preserved as normal and which should be rejected as anomalous. Under limited normal observations, the estimated structure may fail to cover legitimate unseen variations, while directions explaining normal features may also absorb anomaly-induced changes. Neural networks extend data space modeling through learnable nonlinear representations. Autoencoders learn latent representations through input reconstruction, and Deep SVDD learns compact representations of normal observations E et al. (2006); Ruff et al. (2018). Although these methods capture complex data structures, greater representational capacity does not guarantee generalization to unseen feature variations under limited supervision. These limitations highlight the need for more reliable normality modeling under sparse supervision. In particular, preserving

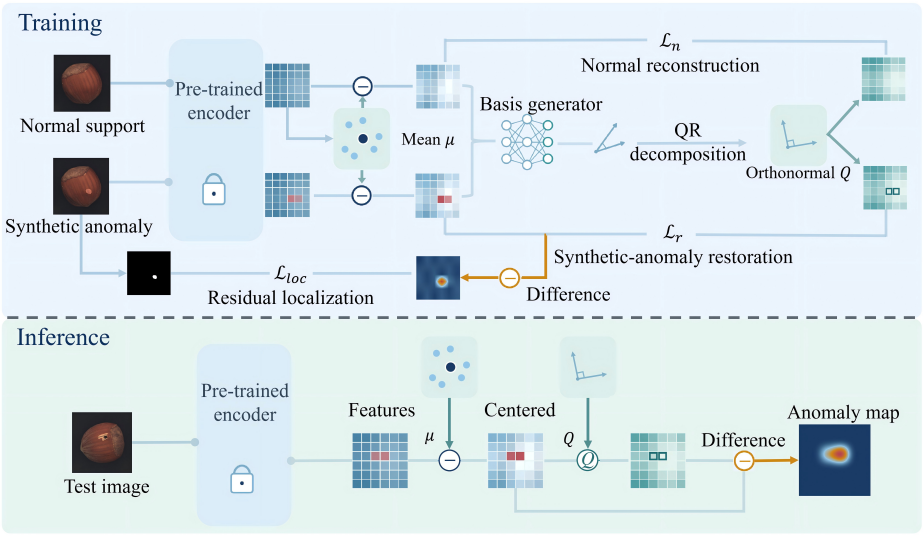

Figure 2: Overview of NormProj. During training, an orthonormal basis is generated from augmented normal support features and optimized through normal reconstruction, paired synthetic-anomaly restoration, and residual localization. During inference, the learned normal representation is used for projection-residual anomaly scoring without the basis generator.

legitimate normal variations while maintaining sensitivity to anomalous changes remains an important challenge in few-shot anomaly detection.

3 Method

3.1 Problem Setup and Motivation

Given a target category, few-shot industrial anomaly detection uses a support set $\mathcal{S} = \{x_j\}_{j=1}^{K}$ of K normal images to detect and localize anomalies in unseen test images. A frozen visual encoder F extracts patch features $F(x) \in \mathbb{R}^{N\times d}$, where N and d denote the number of patches and the feature dimension.

Instead of predicting specific normal features through a nonlinear reconstruction mapping, we seek a category-specific normal subspace that preserves legitimate normal variations while rejecting anomalous changes. Let $Q \in \mathbb{R}^{d\times r}$ be an orthonormal basis with $Q^\top Q = I_r$ and $r < d$, and let $P = QQ^\top$ be its orthogonal projector. For a normal feature z_n, we define $u = z_n - \mu$, where μ is the category-specific normal mean. Normal reconstruction encourages Pu to retain normal information, but does not explicitly prevent anomaly-induced changes from entering the subspace.

To introduce anomaly-excluding supervision, consider a paired synthetic-anomaly feature $\tilde{u} = u + \delta$, where δ denotes the anomaly-induced feature change. Restoring $\tilde{u}$ to its normal counterpart yields

$$\begin{aligned} \|u - P\tilde{u}\|_2^2 &= \|(I-P)u - P\delta\|_2^2 \\ &= \underbrace{\|(I-P)u\|_2^2}_{\text{Normal reconstruction error}} + \underbrace{\|P\delta\|_2^2}_{\text{Retained anomaly energy}} . \end{aligned} \tag{1}$$

Since P is an orthogonal projector, $(I-P)P = 0$, making the two components orthogonal and eliminating the cross term. Minimizing the first term encourages normal preservation, while minimizing the second reduces the anomaly-induced changes retained within the subspace. Since the total energy of each anomaly-induced change is independent of the learned

subspace, reducing its retained energy increases its energy in the orthogonal complement, thereby encouraging anomaly exclusion.

Based on this principle, NormProj learns an anomaly-aware normal subspace through the orthogonal-basis generator and preserve-and-reject learning, as illustrated in Fig. 2.

3.2 Orthogonal Basis Generator

Given V augmented normal support views, the frozen encoder extracts features $Z_v = F(x_v) \in \mathbb{R}^{N\times d}$. We estimate the category-specific normal mean as $\mu = \frac{1}{VN}\sum_{v=1}^{V}\sum_{i=1}^{N} Z_{v,i}$ and keep it fixed during optimization. Each centered feature matrix is denoted by $X_v = Z_v - \mathbf{1}_N\mu^\top$.

To learn normal variation directions, an orthogonal-basis generator g_θ maps the spatial responses of each feature channel into r candidate basis coefficients. The candidate bases are then aggregated across normal views and orthogonalized,

$$B_v = g_\theta(X_v^\top) \in \mathbb{R}^{d\times r}, \quad \bar{B} = \frac{1}{V}\sum_{v=1}^{V} B_v, \quad Q = \mathrm{QR}(\bar{B}). \tag{2}$$

Here, g_θ is a multilayer perceptron, and $\mathrm{QR}(\cdot)$ returns the orthonormal factor of the reduced QR decomposition. The resulting basis satisfies $Q^\top Q = I_r$ and defines a rank-r normal subspace for the target category. As illustrated in Fig. 2, only normal support features enter the basis generator, while synthetic anomalies guide its optimization through the training objectives.

3.3 Preserve-and-Reject Learning

As shown in the training stage of Fig. 2, we adopt Natural Synthetic Anomalies Schlüter et al. (2022) with an additional foreground constraint to construct paired normal and synthetic-anomaly images. Each pair contains a normal image x, its corrupted counterpart $\tilde{x}$, and an anomaly mask M. Their spatially aligned, centered patch features are $u_i = F(x)_i - \mu$ and $\tilde{u}_i = F(\tilde{x})_i - \mu$. Both branches use the learned projector $P = QQ^\top$.

To preserve legitimate normal variations, we minimize the reconstruction error of normal features. However, normal reconstruction alone does not explicitly prevent the subspace from explaining anomalous changes. We therefore introduce paired synthetic-anomaly restoration, which projects corrupted features toward their corresponding clean features. The two objectives are

$$\mathcal{L}_n = \mathbb{E}_i\left[\|u_i - Pu_i\|_2^2\right], \quad \mathcal{L}_r = \mathbb{E}_i\left[\|u_i - P\tilde{u}_i\|_2^2\right]. \tag{3}$$

As established in Eq. equation 1, synthetic-anomaly restoration penalizes normal reconstruction error and anomaly-induced changes retained within the subspace. Together, the two objectives encourage normal preservation and anomaly exclusion.

While these objectives constrain feature-space geometry, they do not directly align projection residuals with anomaly locations. We therefore introduce localization supervision using the synthetic mask. The patch-level residual map is defined as $R_s(i) = \|(I-P)\tilde{u}_i\|_2$. After upsampling and per-image min-max normalization, the resulting map $\hat{R}_s$ is supervised using Dice loss Milletarì et al. (2016),

$$\mathcal{L}_{\mathrm{loc}} = \mathrm{DiceLoss}(\hat{R}_s, M). \tag{4}$$

The overall training objective is

$$\mathcal{L} = \mathcal{L}_n + \mathcal{L}_r + \mathcal{L}_{\mathrm{loc}}. \tag{5}$$

All three objectives optimize the orthogonal-basis generator through Q, while the visual encoder and category-specific normal mean remain fixed.

3.4 Anomaly Scoring and Inference

As illustrated in Fig. 2, the basis generator is discarded after training. Only the category-specific normal mean μ and orthonormal basis Q are retained.

For a test patch feature z, orthogonal projection gives its normal reconstruction $\hat{z} = \mu + QQ^\top(z - \mu)$. The anomaly score is the squared projection residual,

$$s(z) = \|z - \hat{z}\|_2^2 = \|(I - QQ^\top)(z - \mu)\|_2^2. \tag{6}$$

Patch scores are reshaped into an anomaly map, upsampled, and Gaussian-smoothed. The image-level score is the mean of the top one percent of pixel scores.

Projection residuals are invariant to feature changes along the learned normal directions, allowing NormProj to explain unseen normal variations within the subspace without predicting their specific normal counterparts.

4 Experiments

4.1 Experimental Setup

Datasets and Evaluation Metrics. We evaluate NormProj on two widely used industrial anomaly detection benchmarks, MVTec-AD (Bergmann et al., 2019) and VisA (Zou et al., 2022). MVTec-AD contains 15 categories, including 10 object and 5 texture categories, while VisA comprises 12 categories with complex structures, multiple instances, and diverse defect types. Together, they enable evaluation across different normal appearance variations and anomaly patterns. We consider 1-shot, 2-shot, and 4-shot settings, using only the corresponding number of normal support images per category. We report image-level AUROC and AUPR for anomaly detection, and pixel-level AUROC and AU-PRO for anomaly localization. All metrics are reported as category-wise macro averages.

Implementation Details. We employ a frozen DINOv2-B (Oquab et al., 2024) as the visual encoder with an input resolution of 672×672. Patch features from four intermediate layers are averaged to obtain a 48×48 feature grid with 768 channels. For each category, we train an independent orthogonal-basis generator consisting of a two-layer MLP. The normal subspace rank is set to $r = 512$. We employ foreground-constrained Natural Synthetic Anomalies (NSA; (Schlüter et al., 2022)) to generate paired normal and synthetic-anomaly images. The generator is optimized using Adam with a learning rate of 10^{-3}. Training runs for at least 50 and at most 200 epochs, with an early-stopping patience of 10 epochs. We conduct three independent experiments with different random seeds for each few-shot setting and report the mean results.

Comparison Methods. We compare NormProj with representative methods covering different anomaly detection paradigms. SPADE (Cohen & Hoshen, 2020) and PatchCore (Roth et al., 2022) detect anomalies through feature matching against normal references. WinCLIP (Jeong et al., 2023), AnomalyCLIP (Zhou et al., 2024), PromptAD (Li et al., 2024), and AdaptCLIP (Gao et al., 2026) exploit pretrained visual-semantic representations or prompt-based adaptation. InCTRL (Zhu & Pang, 2024) and LogSAD (Zhang et al., 2025) further incorporate contextual information or multimodal reasoning. We also compare with FoundAD (Zhai et al., 2026), which learns a nonlinear projector for normal feature reconstruction, to evaluate the effectiveness of normal subspace learning against nonlinear feature reconstruction. We additionally compare with SubspaceAD (Lendering et al., 2026), which estimates a PCA-based normal subspace from frozen visual features, to evaluate the benefit of anomaly-aware subspace learning over normal-only PCA estimation.

4.2 Quantitative Comparison

Table 1 reports the comparison results across three few-shot settings. NormProj achieves the highest image AUROC and image AUPR on both datasets in the 1-shot setting. It

also obtains the second-highest pixel AUROC on MVTec-AD and ties for the highest AU-PRO on VisA. On MVTec-AD, it consistently achieves the highest image AUROC, image AUPR, and AU-PRO across all settings, reaching 97.56%, 99.03%, and 95.04% in the 4-shot setting, respectively. On VisA, NormProj achieves the highest image AUROC across all three settings and remains competitive in image AUPR and region-level localization. Compared with SubspaceAD, NormProj improves image AUROC by 1.47, 0.57, and 0.33 percentage points on MVTec-AD, and by 2.20, 2.53, and 2.63 points on VisA across the three shot settings, respectively. The corresponding image AUPR gains on VisA are 3.24, 3.62, and 4.17 points. These improvements demonstrate the effectiveness of anomaly-aware normal subspace learning over normal-only PCA estimation. Compared with FoundAD, NormProj also improves image AUROC across all shot settings on both datasets and achieves higher AU-PRO on MVTec-AD. However, SubspaceAD obtains higher pixel AUROC on both datasets, while NormProj matches SubspaceAD in AU-PRO on VisA in the 1-shot setting and slightly exceeds it in the 2-shot and 4-shot settings. Overall, NormProj improves image-level discrimination while maintaining competitive anomaly localization performance.

Table 1: Comparison with state-of-the-art methods on MVTec-AD and VisA under the 1-shot, 2-shot, and 4-shot settings. All values are percentages. Best denotes the best result and Second denotes the second-best result.

		MVTec-AD				VisA			
Shot	Method	I-AUROC	I-AUPR	P-AUROC	AU-PRO	I-AUROC	I-AUPR	P-AUROC	AU-PRO
1-shot	SPADE	81.00	90.60	91.20	83.90	79.50	82.00	95.60	84.10
	PatchCore	83.40	92.20	92.00	79.70	79.90	82.80	95.40	80.50
	WinCLIP	93.10	96.50	95.20	87.10	83.80	85.10	96.40	85.10
	InCTRL	91.30	95.20	94.60	87.80	83.20	84.10	89.00	66.70
	AnomalyCLIP	95.20	97.20	94.60	87.60	87.70	87.70	83.20	90.10
	PromptAD	94.60	97.10	95.90	87.90	86.90	88.40	96.70	85.10
	LogSAD	95.50	97.30	97.00	92.50	89.80	90.30	97.50	88.20
	AdaptCLIP	94.50	97.50	94.30	89.50	90.50	92.30	96.80	90.70
	FoundAD	93.28	96.10	95.99	92.01	89.79	89.20	97.80	93.74
	SubspaceAD	95.37	97.66	97.24	94.25	89.70	89.07	97.52	94.00
	NormProj (Ours)	96.84	98.65	97.01	94.45	91.90	92.31	97.25	94.00
2-shot	SPADE	82.90	91.70	92.00	85.70	80.70	82.30	96.20	85.70
	PatchCore	86.30	93.80	93.30	82.30	81.60	84.80	96.10	82.60
	WinCLIP	94.40	97.00	96.00	88.40	84.60	85.80	96.80	86.20
	InCTRL	91.80	95.50	95.20	88.30	86.30	86.80	89.80	68.10
	AnomalyCLIP	95.40	97.30	94.90	87.80	87.80	89.10	84.50	90.80
	PromptAD	95.70	97.90	96.20	88.50	88.30	90.00	97.10	85.80
	LogSAD	96.30	97.60	97.30	93.10	91.80	92.00	97.80	89.70
	AdaptCLIP	95.70	97.90	94.50	90.30	92.20	93.60	97.10	91.20
	FoundAD	94.95	97.36	96.39	92.88	90.71	88.83	97.97	94.05
	SubspaceAD	96.68	98.23	97.40	94.59	90.72	89.72	97.65	94.43
	NormProj (Ours)	97.25	98.80	97.29	94.87	93.25	93.34	97.41	94.45
4-shot	SPADE	84.80	92.50	92.70	87.00	81.70	83.40	96.60	87.30
	PatchCore	88.80	94.50	94.30	84.30	85.30	87.50	96.80	84.90
	WinCLIP	95.20	97.30	96.20	89.00	87.30	88.80	97.20	87.60
	InCTRL	93.10	96.30	95.80	89.50	87.80	88.00	90.20	68.80
	AnomalyCLIP	96.10	97.80	95.50	88.20	88.80	90.10	85.20	91.40
	PromptAD	96.60	98.50	96.50	90.50	89.10	90.80	97.40	86.20
	LogSAD	96.60	97.70	97.50	93.50	93.20	93.40	98.10	90.50
	AdaptCLIP	96.60	98.40	94.80	90.90	93.10	94.30	97.30	91.70
	FoundAD	94.97	97.38	96.51	92.67	91.08	89.61	98.10	94.39
	SubspaceAD	97.23	98.61	97.51	94.83	91.26	89.90	97.73	94.57
	NormProj (Ours)	97.56	99.03	97.37	95.04	93.89	94.07	97.52	94.62

Note: SubspaceAD and FoundAD were trained using features extracted by the same frozen feature extractor as NormProj.

Deployment Footprint. NormProj employs a frozen DINOv2-B/14 backbone and a category-specific orthogonal-basis generator with 1.44M trainable parameters. After training, the generator is discarded, leaving only the category mean vector and rank-512 orthonormal basis for inference. Excluding the shared backbone, this representation requires 393,984 scalars per category, corresponding to approximately 1.50 MiB in FP32. Compared with the 11.8M-parameter FoundAD predictor, the single-category deployment state is approximately 30.0 times smaller. Assuming FoundAD employs a shared predictor across categories, NormProj requires 4.73M and 5.91M scalars for the 12 VisA and 15 MVTec-AD categories, respectively, corresponding to 2.50 and 2.00 times smaller deployment states. These results demonstrate that NormProj provides compact category-specific normal repre-

sentations while eliminating the need for an additional neural reconstruction network during inference.

4.3 Qualitative comparison

Figure 3 presents qualitative comparisons with FoundAD and SubspaceAD on MVTec-AD and VisA. NormProj produces concentrated responses around annotated defects in pill, zipper, tile, candle, chewinggum, and pcb4, while reducing irrelevant responses in surrounding regions. For normal samples such as leather and macaroni2, NormProj generally suppresses background responses, although some responses remain around object structures. These observations suggest that NormProj suppresses responses to normal variations while retaining defect-related information in the projection residual.

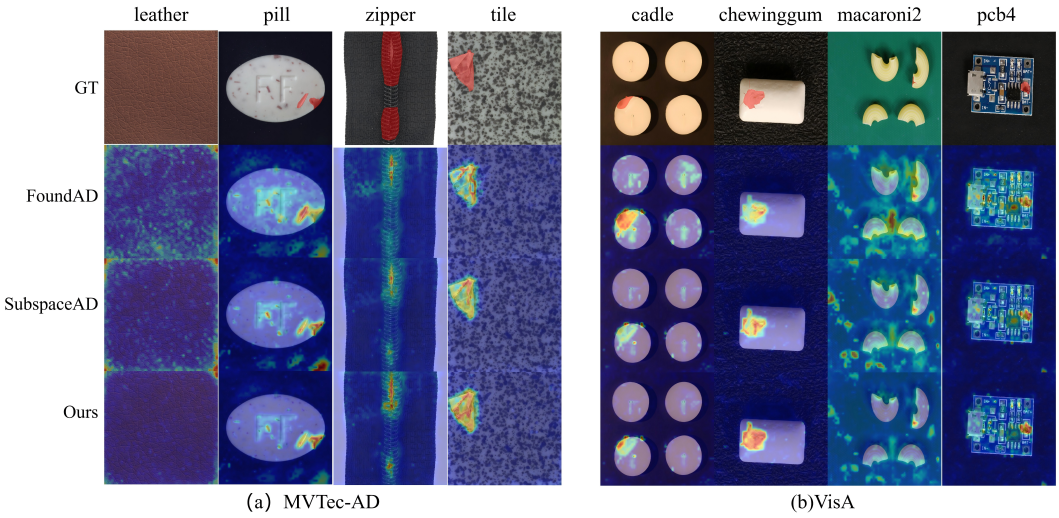

Figure 3: Qualitative comparison of anomaly localization on MVTec-AD and VisA. Red regions in GT indicate ground-truth anomalies, while images without red annotations are normal. Warmer colors in the anomaly maps indicate higher anomaly scores.

4.4 Ablation Studies

Table 2: Ablation of loss components on MVTec-AD under the 1-shot setting. Results are averaged over three random seeds. All values are percentages. The best results are underlined.

Loss configuration	I-AUROC	I-AUPR	P-AUROC	AU-PRO
Normal only	96.07	98.02	<u>97.03</u>	94.08
Normal + restoration	96.51	98.41	97.01	94.06
Full	<u>96.84</u>	<u>98.65</u>	97.01	<u>94.45</u>

Effect of loss components. As shown in Table 2, adding synthetic-anomaly restoration to normal reconstruction improves image AUROC and AUPR by 0.44 and 0.39 percentage points, respectively. Further introducing residual localization improves AU-PRO by 0.39 percentage points. The full objective achieves the highest image AUROC, image AUPR, and AU-PRO, while pixel AUROC remains nearly unchanged across configurations. These results support the complementary roles of normal preservation, anomaly exclusion, and spatial localization in learning the normal subspace.

Effectiveness of the Orthogonal Basis Generator. Table 3 compares different orthogonal basis construction strategies under the MVTec-AD 1-shot setting. Compared with direct basis optimization, NormProj improves image AUROC and AU-PRO by 3.80 and 3.12 percentage

points, respectively, demonstrating the benefit of the proposed basis generator. Compared with normal-only PCA estimation, NormProj further improves image AUROC and AUPR by 1.47 and 0.99 points, while achieving a 0.20-point gain in AU-PRO with a slight decrease in pixel AUROC. These results support the effectiveness of the proposed orthogonal-basis generator and anomaly-aware normal subspace learning.

Table 3: Comparison of orthogonal basis construction strategies on MVTec-AD under the 1-shot setting. Results are averaged over three random seeds. All values are percentages. The best results are underlined.

Method	I-AUROC	I-AUPR	P-AUROC	AU-PRO
Direct-Q	93.04	96.55	94.89	91.33
PCA	95.66	97.73	<u>97.24</u>	94.27
NormProj	<u>96.84</u>	<u>98.65</u>	97.01	<u>94.45</u>

Table 4: Effect of the normal subspace rank on MVTec-AD under the 1-shot setting. All values are percentages. The best results are underlined.

Rank	I-AUROC	I-AUPR	P-AUROC	AU-PRO
128	96.16	98.17	96.59	93.78
256	96.63	98.45	96.96	94.27
384	96.70	98.50	<u>97.02</u>	94.41
512	<u>96.84</u>	<u>98.65</u>	97.01	<u>94.45</u>

Impact of Subspace Capacity. We investigate the effect of subspace rank on anomaly detection and localization. As shown in the rank ablation table, increasing the rank from 128 to 512 improves image AUROC, image AUPR, and AU-PRO, while pixel AUROC peaks at rank 384. Rank 512 achieves the highest image AUROC of 96.84 percent, image AUPR of 98.65 percent, and AU-PRO of 94.45 percent, whereas rank 384 achieves the highest pixel AUROC of 97.02 percent. Pixel AUROC remains nearly unchanged between ranks 384 and 512, suggesting that increasing subspace capacity does not uniformly improve all evaluation metrics.

5 Conclusion

We present NormProj, which advances normal subspace modeling through joint normal preservation and anomaly exclusion. Experiments on MVTec-AD and VisA demonstrate improved image-level anomaly detection and competitive localization performance across few-shot settings. By discarding the basis generator after training, NormProj replaces neural feature reconstruction with compact orthogonal projection, reducing deployment storage and inference-time computation. These findings highlight the potential of anomaly-aware normal subspace learning for few-shot industrial anomaly detection, where preserving legitimate normal variations and maintaining sensitivity to anomalous changes are both essential.

Use of Large Language Models Large language models were used solely for light editing tasks including grammar correction, spelling checks, and minor phrasing improvements to enhance clarity and concision.

References

Principal component analysis. Springer, 14(27):231–246, 1986.

Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation, 15(6):1373—1396, 2003.

Niamh Belton, Misgina Tsighe Hagos, Aonghus Lawlor, and Kathleen M. Curran. Fewsome: One-class few shot anomaly detection with siamese networks. In 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 2978–2987, 2023. doi: 10.1109/CVPRW59228.2023.00299.

Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9584–9592, 2019. doi: 10.1109/CVPR.2019.00982.

Emmanuel J. Candes, L. I. Xiaodong, M. A. Yl, and John Wright. Robust principal component analysis? Journal of the ACM (JACM), 58(3), 2011.

Niv Cohen and Yedid Hoshen. Sub-image anomaly detection with deep pyramid correspondences. ArXiv, abs/2005.02357, 2020. URL https://api.semanticscholar.org/CorpusID:218502727.

Simon Damm, Mike Laszkiewicz, Johannes Lederer, and Asja Fischer. Anomalydino: Boosting patch-based few-shot anomaly detection with dinov2. In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 1319–1329, 2025. doi: 10.1109/WACV61041.2025.00136.

Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: A patch distribution modeling framework for anomaly detection and localization. In Alberto Del Bimbo, Rita Cucchiara, Stan Sclaroff, Giovanni Maria Farinella, Tao Mei, Marco Bertini, Hugo Jair Escalante, and Roberto Vezzani (eds.), Pattern Recognition. ICPR International Workshops and Challenges, pp. 475–489, Cham, 2021. Springer International Publishing. ISBN 978-3-030-68799-1.

G E, Hinton, R R, and Salakhutdinov. Reducing the dimensionality of data with neural networks. Science (New York, N.Y.), 2006.

Zheng Fang, Xiaoyang Wang, Haocheng Li, Jiejie Liu, Qiugui Hu, and Jimin Xiao. Fastrecon: Few-shot industrial anomaly detection via fast feature reconstruction. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 17435–17444, 2023. doi: 10.1109/ICCV51070.2023.01603.

Bin-Bin Gao, Yue Zhou, Jiangtao Yan, Yuezhi Cai, Weixi Zhang, Meng Wang, Jun Liu, Yong Liu, Lei Wang, and Chengjie Wang. AdaptCLIP: Adapting clip for universal visual anomaly detection. In Proceedings of the AAAI Conference on Artificial Intelligence, pp. 4095–4103, 2026. doi: 10.1609/aaai.v40i6.42404. URL https://ojs.aaai.org/index.php/AAAI/article/view/42404.

Zhaopeng Gu, Bingke Zhu, Guibo Zhu, Yingying Chen, Ming Tang, and Jinqiao Wang. Univad: A training-free unified model for few-shot visual anomaly detection. In 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 15194–15203, 2025. doi: 10.1109/CVPR52734.2025.01415.

Chaoqin Huang, Haoyan Guan, Aofan Jiang, Ya Zhang, Michael Spratling, and Yan-Feng Wang. Registration based few-shot anomaly detection. In Shai Avidan, Gabriel Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), Computer Vision – ECCV 2022, pp. 303–319, Cham, 2022. Springer Nature Switzerland. ISBN 978-3-031-20053-3.

Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, and Onkar Dabeer. Winclip: Zero-/few-shot anomaly classification and segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 19606–19616, June 2023.

D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755):788–791, 1999.

Camile Lendering, Erkut Akdag, and Egor Bondarau. Subspacead: Training-free few-shot anomaly detection via subspace modeling. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 28557–28566, June 2026.

Xiaofan Li, Zhizhong Zhang, Xin Tan, Chengwei Chen, Yanyun Qu, Yuan Xie, and Lizhuang Ma. Promptad: Learning prompts with only normal samples for few-shot anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16838–16848, June 2024.

Wenxi Lv, Qinliang Su, and Wenchao Xu. One-for-all few-shot anomaly detection via instance-induced prompt learning. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), International Conference on Learning Representations, volume 2025, pp. 63657–63677, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/9f7f2f57d8eaf44b2f09020f64ff6d96-Paper-Conference.pdf.

Michael, E., Tipping, Christopher, M., and Bishop. Probabilistic principal component analysis. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 1999.

Fausto Milletarì, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571, 2016. URL https://api.semanticscholar.org/CorpusID:206429151.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy V. Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel HAZIZA, Francisco Massa, Alaaeldin El-Nouby, Mido Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Herve Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. DINOv2: Learning robust visual features without supervision. Transactions on Machine Learning Research, 2024. ISSN 2835-8856. URL https://openreview.net/forum?id=a68SUt6zFt. Featured Certification.

Zhen Qu, Xian Tao, Xinyi Gong, ShiChen Qu, Xiaopei Zhang, Xingang Wang, Fei Shen, Zhengtao Zhang, Mukesh Prasad, and Guiguang Ding. Dictas: A framework for class-generalizable few-shot anomaly segmentation via dictionary lookup. In 2025 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 20519–20528, 2025. doi: 10.1109/ICCV51701.2025.01908.

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

St Roweis and Lk Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500):2323–2326, 2000.

Lukas Ruff, Robert Vandermeulen, Nico Goernitz, Lucas Deecke, Shoaib Ahmed Siddiqui, Alexander Binder, Emmanuel Müller, and Marius Kloft. Deep one-class classification. In Jennifer Dy and Andreas Krause (eds.), Proceedings of the 35th International Conference on Machine Learning, volume 80 of Proceedings of Machine Learning Research, pp. 4393–4402. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/ruff18a.html.

Hannah M. Schlüter, Jeremy Tan, Benjamin Hou, and Bernhard Kainz. Natural synthetic anomalies for self-supervised anomaly detection and localization. In Shai Avidan, Gabriel

Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), Computer Vision – ECCV 2022, pp. 474–489, Cham, 2022. Springer Nature Switzerland. ISBN 978-3-031-19821-2.

B Schölkopf, A Smola, and K Müller. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation, 10(5):1299–1319, 1998.

Joshua B. Tenenbaum, Vin De Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. Science, 290(5500):2319–2323, 2000.

Guangyao Zhai, Yue Zhou, Xinyan Deng, Lars Heckler-Kram, Nassir Navab, and Benjamin Busam. Foundation visual encoders are secretly few-shot anomaly detectors. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), International Conference on Learning Representations, volume 2026, pp. 97353–97376, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/9e47a0bc530cc88b09b7670d2c130a29-Paper-Conference.pdf.

Jinjin Zhang, Guodong Wang, Yizhou Jin, and Di Huang. Towards training-free anomaly detection with vision and language foundation models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 15204–15213, June 2025.

Wenchao Zhang, Li Wang, and Weichen Li. Rethinking the prior guided vision transformer for industrial anomaly class discovery. Applied Soft Computing, 203:116078, 2026. ISSN 1568-4946. doi: https://doi.org/10.1016/j.asoc.2026.116078. URL https://www.sciencedirect.com/science/article/pii/S1568494626015267.

Qihang Zhou, Guansong Pang, Yu Tian, Shibo He, and Jiming Chen. Anomalyclip: Object-agnostic prompt learning for zero-shot anomaly detection. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, and Y. Sun (eds.), International Conference on Learning Representations, volume 2024, pp. 49705–49737, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/d7b50b8ac2c781a12f26155f48310d8d-Paper-Conference.pdf.

Jiawen Zhu and Guansong Pang. Toward generalist anomaly detection via in-context residual learning with few-shot sample prompts. In 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 17826–17836, 2024. doi: 10.1109/CVPR52733.2024.01688.

Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. Spot-the-difference self-supervised pre-training for anomaly detection and segmentation. In Shai Avidan, Gabriel Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (eds.), Computer Vision – ECCV 2022, pp. 392–408, Cham, 2022. Springer Nature Switzerland. ISBN 978-3-031-20056-4.