PRETRAIN ONCE, DEPLOY MANY TIMES: FREQUENCY-AWARE MEMORY FOR TRAINING-FREE ANOMALY DETECTION ON UNSEEN MACHINES

Anonymous authors
Paper under double-blind review

ABSTRACT

Fault diagnosis foundation models promise reusable vibration representations, yet deployment still commonly requires target-specific training for each new machine. We introduce a pretrain-once, deploy-many-times framework: a vibration encoder is pretrained once on a heterogeneous cross-source corpus, frozen globally, and reused on unseen targets without target-domain gradient updates, while each target defines normality locally from a small set of normal references. FAMB (Frequency-Aware Memory Bank) instantiates this local detector through frequency-domain structure: it anchors matching to local Mel-frequency neighbourhoods and uses second-order latent frequency-shape residuals for bounded score enhancement. Same-sensor matching and unrestricted time positions preserve meaningful frequency comparisons. On four datasets excluded from pretraining, FAMB with one normal reference sample reaches 96.25%/96.92% macro-average AUROC/AUPR. On 16 liquid rocket engines, 3-fold device-level cross-validation keeps reference and test devices disjoint. With 100 samples from one normal engine, FAMB reaches 96.49%/96.88% AUROC/AUPR (across-fold standard deviation 1.98%/2.14%), exceeding the strongest baseline by 3.80 and 4.28 percentage points. These results support pretrain-once, deploy-many-times fault diagnosis through globally reusable vibration representations and locally defined normality.

1 INTRODUCTION

Industrial fault diagnosis has traditionally relied on machine- or domain-specific models, retrained whenever the sensing environment changes. The promise of foundation models is different: representation learning should be amortised across deployments, so that one encoder trained once can serve many future machines (Eldele et al., 2025; Sehri et al., 2025). Recent work explores generative diagnosis, unified modules and self-supervised vibration Transformers (Wang et al., 2025; Li et al., 2025; Mannone et al., 2026), but this pretrain-once, reuse-many-times capability remains underexplored for vibration: a transferable encoder does not by itself define what is normal on an unseen machine, so deployment conventionally still requires target-specific training.

We adopt a separation of responsibilities: representation learning is global, whereas normality modelling is local. A single encoder is pretrained on a heterogeneous cross-source vibration corpus and frozen permanently; each target setting instantiates its own detector from a small normal reference set, reducing deployment to memory construction and inference. Training-free refers to this local stage: anomalous examples and gradient updates are unnecessary. Crucially, all evaluation datasets and physical test machines are excluded from pretraining.

A reusable encoder alone does not specify what is normal for a newly encountered machine, and that normality must be compared in a meaningful frequency coordinate system. Image-based memory methods (Roth et al., 2022; Defard et al., 2021) provide a retrieval framework, but vibration spectrograms need band-aware comparisons. FAMB anchors retrieval to neighbouring Mel-frequency rows and adds second-order variation of the latent frequency profile. Sensor identity fixes sensing

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.

position, whereas free time matching accommodates windowing phase; together they preserve local frequency correspondence without exact temporal alignment.

Our contributions are threefold. (i) Pretrain-once, deploy-many-times fault diagnosis: a deployment regime separating global vibration representation learning from local normality modelling, where one encoder is reused on targets excluded from pretraining with no target-domain gradient updates. (ii) Frequency-aware training-free adaptation: FAMB instantiates machine-specific normality from normal references alone, through sensor-conditioned local frequency retrieval, unrestricted temporal matching, and bounded second-order latent frequency-shape enhancement. (iii) Deployment-oriented evaluation: under a shared frozen backbone we test unseen-source transfer, few-shot references and device-disjoint engines, including normality transfer from a single reference engine. Figure 1 summarises this regime.

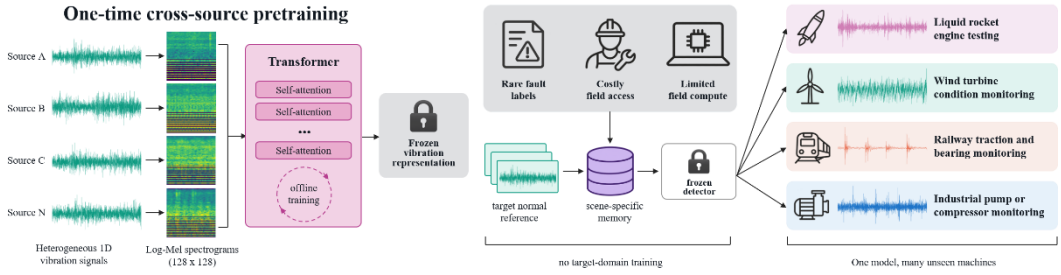


Figure 1: Pretrain once, deploy many times. One cross-source corpus is used once to learn a reusable vibration encoder; the frozen encoder is then shared across deployment sites, each instantiating its own detector from local normal references. Remote engine test stands, wind farms, railway traction and pump stations illustrate settings where retraining at every site is costly.

2 Related Work and Problem Setting

Representation pretraining. DINO, MAE, I-JEPA, MoCo v3, BYOL, SwAV and VICReg represent major self-supervised visual learners (Caron et al., 2021; He et al., 2022; Assran et al., 2023; Chen et al., 2021; Grill et al., 2020; Caron et al., 2020; Bardes et al., 2022), while Transformers support visual and time-series representations (Dosovitskiy et al., 2021; Wu et al., 2023; Nie et al., 2023). Their evaluations use different modalities and protocols. We test whether cross-source vibration pretraining improves frozen features for unseen sources while holding the input, backbone capacity and downstream pipeline fixed.

Vibration representations and multi-condition diagnosis. Work on transferable diagnosis spans vibration resources, machinery foundation models, frequency priors, domain generalisation, meta-learning and continual learning (Eldele et al., 2025; Sehri et al., 2025; Wang et al., 2022; Zhao et al., 2020; Han et al., 2025; Xiao et al., 2025; Ren et al., 2024; Liu et al., 2026; Chen et al., 2026; Jung et al., 2023; Zhang et al., 2018; Li et al., 2023; Lian et al., 2025; Lei et al., 2020; Yu et al., 2022; Lai et al., 2024; Singh, 2025; Li et al., 2024). Because these methods differ in corpus scale, operating conditions and splits, we enforce explicit boundaries: every target dataset is excluded from pretraining, and deployment uses its own normal references, sensor identities and frequency coordinates with a shared frozen encoder.

Time-series foundation models and anomaly detection. Anomaly Transformer detects anomalies through association discrepancy (Xu et al., 2022). Later work studies foundation models, cross-machine transfer, zero-shot detection and online memory (Feng et al., 2022; Xu et al., 2024; Fang et al., 2024; Hermary et al., 2026; He et al., 2026; Chen et al., 2024; Khan et al., 2026; Li et al., 2022; Wenig et al., 2024; Wen & Chen, 2026; Fu et al., 2026; Xia et al., 2026; Zhao et al., 2019). We evaluate normal-reference adaptation through detection heads under one implementation and data protocol.

Normality modelling and the closest work. PatchCore stores local normal features, PaDiM models position-dependent distributions, Deep SVDD learns a one-class boundary, and RD4AD/Dinomaly reconstruct normal features (Roth et al., 2022; Defard et al., 2021; Ruff et al., 2018; Deng & Li, 2022; Guo et al., 2025). Recent visual detectors also use pretrained features with few normal references (Xie et al., 2024; Gao et al., 2025; Fang et al., 2025; Yao et al., 2025; Lin et al., 2026; Cheng et al., 2025; Batzner et al., 2024; Gao & Wang, 2026). RAD (Zhang et al., 2026) is closest: it enables training-free detection with frozen features. FAMB extends this retrieval setting: instead of one isotropic neighbourhood, it anchors frequency, frees time, requires identical sensors and adds second-order latent frequency-shape residuals.

Problem setting. Given a pretraining corpus D_{pre} and target normal references D_N^T, we learn an encoder f_θ on D_{pre}, freeze θ, and build a target memory M_T and normal statistics from D_N^T. Each target setting defines its own sensor identities and memory. Target datasets are excluded from D_{pre}; mirrored or duplicate records are removed by record identity, and raw records are split before windowing. Target anomalies and test windows enter neither memory construction nor parameter estimation. We evaluate two regimes: *unseen-source deployment*, where the target is absent from pretraining and only its normal data instantiate the detector (public datasets), and *cross-device transfer*, where references and test queries come from physically distinct units (engine fleet).

3 METHOD

3.1 CROSS-SOURCE SELF-DISTILLATION PRETRAINING AND TARGET INPUTS

Figure 2 outlines the pipeline. The cross-source corpus D_{pre} comprises 38 public vibration sources (Appendix A) covering bearings, gears and gearboxes, rotor dynamics, electric motors and rotating-machinery rigs; it excludes all four public evaluation datasets and all engine test devices. Each vibration channel is independently windowed and converted to a log-Mel spectrogram $M \in \mathbb{R}^{128\times128}$ (Mannone et al., 2026). The encoder processes one channel at a time, and memory entries retain sensor identity. The input transform is

$$M \;=\; \mathrm{PowerToDB}\Big(\Phi_{\mathrm{mel}}(FS)\cdot|\mathrm{STFT}(\widetilde{x})|^2\Big) \tag{1}$$

Here $\widetilde{x}$ denotes the z-scored vibration window, and $\Phi_{\mathrm{mel}}(FS)$ the Mel filter bank constructed for the sampling frequency FS of that source (Table 1). The 128×128 representation gives every source a common encoder interface, but physical frequency coverage still depends on sampling rate and preprocessing. References and queries within a target setting share the transform, making local Mel coordinates comparable without implying a shared mechanical order across datasets.

The backbone uses patch projection, positional encoding and a Transformer (Dosovitskiy et al., 2021). DINO-style multi-view self-distillation (Caron et al., 2021) trains a student on all views and an exponential-moving-average teacher on global views, with $\xi \leftarrow m\xi + (1-m)\theta$. We export the teacher backbone and freeze it for downstream evaluation. Patch tokens from layers 1, 3, 5 and 7 populate separate memories. Equations (2)–(5) are evaluated per layer, after which s_{base} and a_2 are each averaged equally across the four layers before Eq. (6).

Channel-level and per-window z-scores plus window-maximum PowerToDB remove absolute amplitude comparability. "Frequency shape" therefore denotes variation of normalised latent features across discrete frequency rows, modeled by differences rather than absolute energy.

3.2 FAMB: CONDITIONED RETRIEVAL AND SECOND-ORDER FREQUENCY-SHAPE ENHANCEMENT

Design principle. FAMB instantiates the target detector: it converts a small normal reference set into machine-specific normality while leaving the shared encoder unchanged. It models normality through local frequency correspondence and latent frequency shape. Same-sensor matching preserves sensing semantics; unrestricted time matching accommodates arbitrary windowing phase.

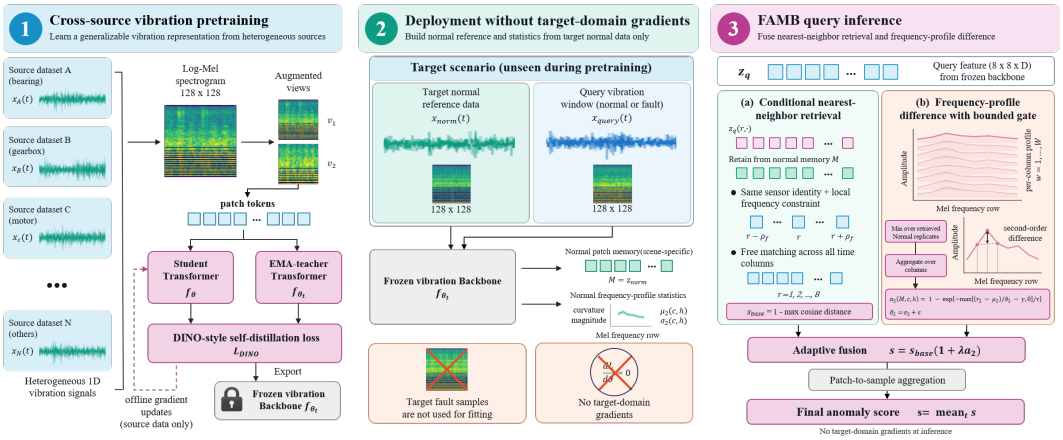


Figure 2: Pretrain once, deploy without target-domain gradients. Cross-source data trains the encoder (left). Target normal references supply memory and frequency statistics (middle). Conditioned retrieval and bounded shape enhancement score queries (right).

Candidate conditioning. For normal reference i, the frozen encoder produces L2-normalised patch features $u_i[n]$. Each entry retains its sensor c_n, Mel-frequency row h_n and time column w_n. These entries form the target memory M_T (Roth et al., 2022; Zhang et al., 2026). For query M, we first rank reference windows by cosine similarity of global descriptors and retain the $\text{top}_k = 8$ windows, or all windows if fewer are available. This set, $T(M)$, limits retrieval to globally similar references. The descriptor is the L2-normalised mean of the layer-7 patch tokens, the deepest of the four layers that populate the memory; the backbone has no CLS token. For query patch t with L2-normalised feature $q(M,t)$, candidates must then share its sensor and lie within a local frequency neighbourhood. Time position is unrestricted:

$$C(M,t) \;=\; \{u_i[n] \in M_T \;:\; i \in T(M),\; c_n = c_t,\; |h_n - h_t| \le \rho_f\} \tag{2}$$

Free time matching accommodates arbitrary window boundaries (Section 4.5). Coordinate constraints also reduce retrieval: with roughly uniform sensor and frequency coverage, candidates occupy about $(2\rho_f + 1)/(N_{\text{sen}} \cdot H)$ of the memory retained by $T(M)$, where N_{sen} is the number of sensors, that is, of recorded channels, and H the number of patch frequency rows. The 16×16 patch size gives $H = 8$ and 16 Mel bins per row. With $\rho_f = 1$, retrieval admits the query row and its two neighbours. The base score is the nearest-candidate cosine distance (Zhang et al., 2026):

$$s_{\text{base}}(M,t) \;=\; 1 \;-\; \max_{u \in C(M,t)} \langle q(M,t),\, u \rangle \tag{3}$$

Second-order frequency-shape enhancement. Retrieval measures local feature coverage, but it does not directly compare how features vary across frequency. For each reference window j and time column w, we renormalise the tokens of sensor c along frequency to form a latent frequency profile $P^{(j,w)}(c,\cdot)$ and take second-order differences along frequency, storing their magnitudes. Time carries no fixed correspondence, so the W columns of a window serve as replicates of the same profile, a proxy for short-time spectral fluctuation: the replicate set is $\{(j,w)\}$ rather than $\{j\}$, of size $n \cdot W$ instead of n. Differences use valid rows only: boundary rows without two neighbours receive zero residual and zero enhancement, retaining their base scores. The residual of a query column w is its distance to the closest replicate among the selected references:

$$r_2^{(w)}(M,c,h) \;=\; \min_{j \in T(M),\, w'} \left| \left\| D_2\left(P_M^{(w)}\right)(c,h) \right\|_2 \;-\; \left\| D_2\left(P^{(j,w')}\right)(c,h) \right\|_2 \right| \tag{4}$$

where $T(M)$ contains the selected normal references. Curvature magnitudes capture local bending without requiring agreement in every feature direction, and r_2 averages residuals over time columns.

Normal scale uses leave-one-replicate distances: each normal replicate is scored against the nearest replicate of another window within its own top_k set, matching the test-time pool, or, for $n = 1$, the nearest other column of the same window. Thus μ_2 and σ_2 remain estimable without a separate calibration set. We estimate them per sensor, row and layer, shrink them toward pooled values ($\kappa = 8$), floor the pooled σ at 5% of its mean and add $\varepsilon = 10^{-12}$ for stability; we write $\widetilde{\sigma}_2$ for σ_2 after shrinkage, flooring and stabilisation, and it is this shrunken estimate that standardises the residual below. This keeps $\widetilde{\sigma}_2$ positive and a_2 graded as the reference budget changes. The standardised residual above margin γ is mapped through temperature $\tau > 0$ to $[0, 1]$:

$$a_2(M, c, h) \;=\; 1 \;-\; \exp\left\{-\max\left[\left(r_2 - \mu_2\right)/\widetilde{\sigma}_2 - \gamma,\; 0\right]/\tau\right\} \tag{5}$$

Equation (5) applies at every $n \geq 1$, where n counts reference windows and $W = 8$ (Appendix B).

Bounded multiplicative fusion. With $\lambda \geq 0$, the coefficient for a patch's sensor and frequency row enhances its nearest-neighbour score. The window score averages these patch scores:

$$s(M, t) \;=\; s_{\mathrm{base}}(M, t)\left[1 + \lambda\, a_2(M, c_t, h_t)\right], \qquad S(M) \;=\; \mathrm{mean}_t\; s(M, t) \tag{6}$$

Because $a_2 \in [0, 1]$ and $\lambda \geq 0$, $s_{\mathrm{base}} \leq s \leq (1 + \lambda) s_{\mathrm{base}}$. The shape branch amplifies existing retrieval evidence by a factor bounded by $1 + \lambda$; a zero base score remains zero. For multi-sensor records, each channel is scored independently and the sample score is the channel mean; channel-max pooling was less accurate and less stable across folds in our tests.

4 Experiments

We evaluate pretrain-once, deploy-many-times along a progressively harder deployment ladder: unseen public datasets (Section 4.2), then at most four normal references (Section 4.3), then normal references and test samples drawn from physically disjoint engines (Section 4.4). Section 4.5 isolates the components that enable this deployment.

4.1 Datasets, splits and evaluation setting

Table 1 summarises JNU (Li et al., 2013), XJTU-G (Liang et al., 2019), SEU-G (Shao et al., 2019) and HUST (Zhao et al., 2024), all excluded from pretraining. Each provides 400 normal references spanning its defined operating conditions and a fixed test set of 200 normal and 200 anomalous samples. Records are split before windowing to prevent segment overlap across sets. The engine data (Lv et al., 2021) contains 12 normal and 4 anomalous machines. In 3-fold device-level cross-validation, each fold uses 8 normal engines as references and 4 as unseen normal test devices. All 4 anomalous engines are test-only in every fold. Each normal engine is held out once, with no device crossing sets within a fold. Engine results report the fold mean and standard deviation.

Table 1: Evaluation data and sample splits. FS is the raw sampling frequency. The public target datasets did not take part in pretraining, and their complete normal reference pool covers all operating conditions. The engine row gives the per-fold reference budget in samples; its test set is fixed at 400 normal + 400 anomalous.

Dataset	FS	Channels	Normal ref.	Normal test	Anomalous test
JNU	50,000 Hz	1	400	200	200
XJTU-G	20,480 Hz	2	400	200	200
SEU-G	12,000 Hz	8	400	200	200
HUST	25,600 Hz	3	400	200	200
Liquid rocket engine	25,600 Hz	11	800 / 400 / 200 / 100	400	400

Engine vibration is sampled at 25.6 kHz. The 11 channels correspond to the same physical sensor positions on every engine: thrust chamber (Ch0–Ch2), gas generator (Ch3–Ch5), oxidiser pump

(Ch6–Ch8), oxidiser pre-pump (Ch9) and second-stage fuel pump (Ch10) (Lv et al., 2021). This correspondence supports same-sensor retrieval across engines. Each machine provides 100 samples. Reference budgets use 8, 4, 2 or 1 of the fold's reference engines, giving 800, 400, 200 or 100 normal samples; the test set remains fixed.

We report threshold-free AUROC and AUPR (%), macro-averaged equally across public datasets; AUPR reflects false-alarm trade-offs (Davis & Goadrich, 2006). Test data never enters memory, parameter estimation or threshold selection. Figure 3 and Tables C1–C2 use five noisy test repetitions (SNR = 5.0 dB); Figure 4 and Table 3 use fixed no-noise engine folds, and Section 4.3 uses 20 fresh reference draws. FPR95 is reported only in Section 4.5. "-style" heads reimplement the cited mechanism under the shared interface.

4.2 Cross-source representation reuse and detection heads

Figure 3 separates the roles of pretraining and detection; Appendix C (Tables C1–C2) gives full per-dataset results. Panel (a) fixes FAMB and varies the frozen representation. Every single-source baseline uses our ViT backbone and is pretrained from scratch for 100 epochs, until its loss converges, on the normal-train split of the evaluation dataset on which it is tested (400 normal samples; its test windows take no part in pretraining), whereas our encoder is pretrained once on Appendix A and never observes any of them. Despite this exclusion, our cross-source backbone leads on every dataset at 99.65%/99.66% macro AUROC/AUPR, against 98.51%/98.56% for the strongest single-source baseline.

The DINO comparison fixes the self-distillation objective, input representation, backbone capacity and downstream detector, leaving the scope of representation learning as the only difference: target-specific versus cross-source pretraining. It accounts for a gain of 1.55 AUROC points, at a comparatively low per-dataset variance across the representations evaluated here (Table C1).

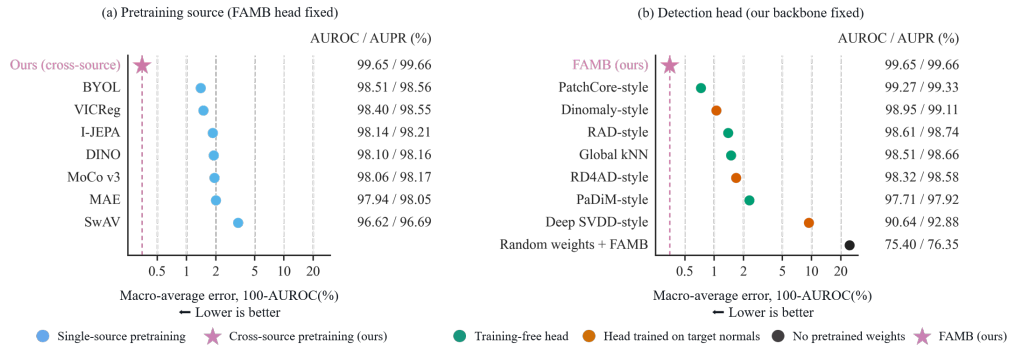


Figure 3: Macro-average discrimination error (100−AUROC, %) on the four unseen datasets, logarithmic axis, lower is better. **(a)** Pretraining sources with the FAMB head fixed. **(b)** Detection heads with our backbone fixed; "-style" names denote our unified re-implementations. Test windows carry Gaussian noise at SNR = 5.0 dB; per-dataset values are in Tables C1–C2.

Panel (b) fixes our backbone and varies the detection head. FAMB reaches 99.65%/99.66% with no target-domain optimisation, exceeding the best trained head (Dinomaly-style, +0.70/0.55) and the best training-free head (PatchCore-style, +0.38/0.33); global retrieval and fixed-position modelling trail by 1.1–1.9 AUROC points. This full-reference setting is near saturation, so it mainly establishes the upper regime: local adaptation matters more as references become scarce and must transfer across devices. Replacing the pretrained backbone with random weights drops macro AUROC to 75.40%, confirming that both components matter. In t-SNE projections of HUST test windows (Figure 5), our encoder separates the two health states more clearly than the seven baselines, although only they were pretrained on HUST.

4.3 FEW-SHOT NORMAL REFERENCES

A deploy-many-times system is useful only if instantiating a new target does not require rebuilding a machine-specific dataset. One reference sample is a raw window, yielding $n\times$ channels spectrograms for n samples; we draw 1, 2 or 4 samples from each dataset's 400-sample normal pool, hold testing fixed and repeat 20 times, updating no parameter. The full pool spans all defined operating conditions; a few-shot draw need not. Table 2 reports macro-averages; Table C3 gives per-dataset means and standard deviations.

Table 2: Macro-average AUROC / AUPR (%) with $n = 1$, 2 and 4 normal references, averaged over 20 repetitions with fresh reference draws and no added noise. Budgets count raw normal windows; patches from one window are not additional references. Test windows are excluded from all fitted statistics.

Detection head	$n=1$ AUROC	$n=1$ AUPR	$n=2$ AUROC	$n=2$ AUPR	$n=4$ AUROC	$n=4$ AUPR
PatchCore-style	92.45	93.90	96.04	96.75	97.22	97.57
RAD-style	93.97	94.90	95.80	96.55	97.08	97.40
Dinomaly-style	91.98	93.72	94.70	96.01	94.91	95.98
FAMB	**96.25**	**96.92**	**97.50**	**97.98**	**98.40**	**98.65**

FAMB leads at all three budgets, with 96.25%/96.92%, 97.50%/97.98% and 98.40%/98.65% AUROC/AUPR. At $n = 1$ the shape branch stays calibrated: the $W = 8$ time columns of the single reference window act as replicates that proxy short-time spectral fluctuation (Section 3.2), so μ_2 and σ_2 are estimated from W leave-one-column distances instead of collapsing to $\widetilde{\sigma}_2 = \varepsilon$. FAMB exceeds RAD-style (93.97%/94.90%), the strongest baseline at this budget, by 2.28/2.02 points. At $n = 2$ and $n = 4$, FAMB exceeds the strongest baseline, PatchCore-style, by 1.46/1.23 and 1.18/1.08 points.

Reference-draw variability is largest at $n = 1$ for every baseline, and FAMB's spread decreases with n on JNU, XJTU-G and SEU-G. The one exception is HUST, where FAMB's spread is flat rather than decreasing between $n = 1$ and $n = 4$ (1.25 and 1.29 AUROC points; 0.91 and 0.95 AUPR). At $n = 1$, where a single draw matters most, FAMB's standard deviation is no higher than any baseline's on any dataset (Table C3). This lower variability is consistent with conditioning restricting retrieval to comparable sensor and frequency coordinates.

These experiments isolate reference scarcity, but references and test windows may still originate from the same machine; Section 4.4 therefore imposes the stricter requirement that normality transfer to unseen engines.

4.4 CROSS-DEVICE DEPLOYMENT: LIQUID ROCKET ENGINES

The engine experiment is our strictest test of deploy-many-times: the pretrained encoder stays frozen, normality is instantiated from reference engines, and evaluation runs on physically disjoint engines that contribute neither pretraining data nor memory samples. Each fold follows the device-disjoint split of Section 4.1. Steady-state windows inherit whole-engine health labels; metrics therefore measure health-state discrimination after start-up. Figure 4 shows the split and reference-budget curves; Appendix D provides the engine schematic (Figure 6) and the numerical results (Table D1).

FAMB has the highest mean AUROC at all four reference budgets (99.59%/99.51%, 98.96%/99.04%, 97.57%/98.04% and 96.49%/96.88% AUROC/AUPR), and at the two smallest budgets it also has the lowest across-fold variance. As the budget falls from eight engines to one, it loses only 3.10 AUROC and 2.63 AUPR points, the smallest decrease among the evaluated heads; PatchCore-style loses 6.57/6.78 points and Dinomaly-style 27.66/26.58 (Table D1).

Its advantage grows as deployment becomes more reference-constrained: the AUROC lead over PatchCore-style widens from 0.33 points with eight reference engines to 3.80 with a single one (4.28 AUPR). Frequency-aware local adaptation therefore matters most precisely when machine-specific normal data are scarce.

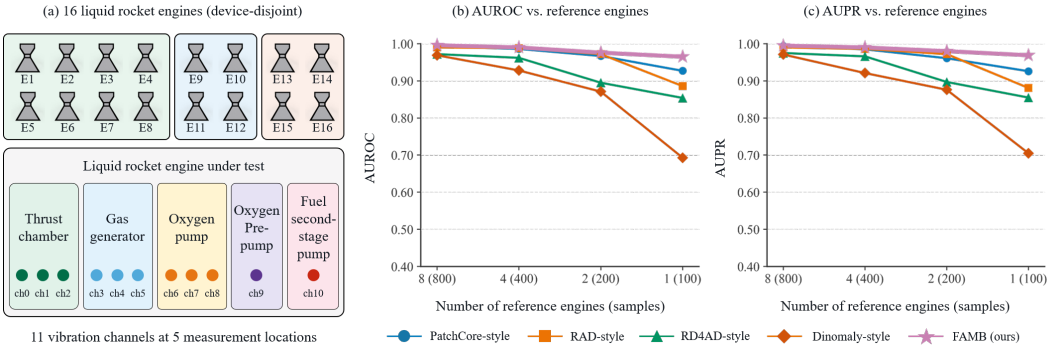


Figure 4: Cross-device engine evaluation. (a) Sensor layout and 3-fold split: each fold uses 8 normal reference engines, 4 unseen normal test engines and all 4 anomalous test engines, with no device crossing sets. (b–c) AUROC and AUPR across reference-engine budgets, shown as fold mean ± standard deviation without added noise. Table D1 gives the numerical values.

Why cross-device is harder. Disjoint devices add machine-to-machine variation to limited reference coverage; sensor and local-frequency constraints reject nearby features from the wrong sensing position or band, which plausibly explains FAMB's engine results. Public and engine data also differ in sampling rate, channel count and composition, and reducing the engine budget changes both sample count and device composition, so neither contrast isolates device separation alone.

Table 3: Structural ablation of FAMB on the engine data (3-fold device-level cross-validation, no added noise). Higher AUROC and lower FPR95 are better; the relative error reduction, computed from the 3-fold means, is the fraction of error removed when that configuration is switched to the full method. Four pairs differ in a single factor: isotropic vs. frequency-anchored (time constraint), frequency-anchored vs. full (shape branch), and full without sensor constraint or without frequency anchoring (ρ_f spanning all H rows) vs. full. The unconditioned row also disables frequency anchoring and the global pre-filter. Cells report the 3-fold mean ± standard deviation.

Configuration	Frequency anchoring	Time axis	Sensor constraint	Second-order shape branch	AUROC %	FPR95 %	Relative error reduction
Unconditioned (match all)	no	free	no	no	96.28 ± 1.01	6.75 ± 0.54	89.0%
Isotropic conditioning	yes	constrained	yes	no	98.01 ± 0.76	5.75 ± 0.22	79.4%
Frequency-anchored + time-free	yes	free	yes	no	98.92 ± 0.72	5.00 ± 0.11	62.0%
Full method without sensor constraint	yes	free	no	yes	99.07 ± 0.64	5.00 ± 0.13	55.9%
Full method without frequency anchoring	no	free	yes	yes	98.95 ± 0.68	4.85 ± 0.15	61.0%
Full method (FAMB)	**yes**	**free**	**yes**	**yes**	**99.59 ± 0.54**	**3.90 ± 0.09**	**—**

4.5 STRUCTURAL ABLATION

We test retrieval constraints and frequency-shape enhancement on the engine data, fixing the 3-fold split, backbone and full 8-engine reference budget; only scoring changes. Table 3 reports AUROC, FPR95 and the relative reduction in $1-\text{AUROC}$, which resolves differences near saturation. FPR95 is read from the test ROC curve for evaluation only and does not set a deployment threshold.

The unconditioned variant, which disables every retrieval constraint and the global pre-filter at once, gives 96.28% AUROC and 6.75% FPR95; it is a weak reference point, not a one-factor ablation or a RAD-style configuration, and the step to 98.01% measures the joint effect of four restored components. Four contrasts isolate a single factor. (i) Temporal tolerance: the isotropic and frequency-anchored rows share the sensor constraint, the radius ρ_f and the pre-filter and differ only in whether time is constrained, and freeing time gives 98.92%, removing 46% of the error remaining at that point. (ii) The shape branch: the frequency-anchored row and the full method differ only in the shape branch, which raises AUROC from 98.92% to 99.59% and, more decisively given the across-fold spread, lowers FPR95 from $5.00 \pm 0.11\%$ to $3.90 \pm 0.09\%$. (iii) Sensor conditioning: removing only the sensor constraint from the full method shows that it is worth 0.52 AUROC points on top of everything else. (iv) Frequency anchoring: letting retrieval span all H rows while keeping the sensor constraint, free time, the pre-filter and the shape branch lowers AUROC to 98.95% and raises FPR95 to 4.85%, so local Mel-row localisation contributes beyond global candidate filtering. In absolute terms the four single-factor gains are 0.91 AUROC points for temporal tolerance, 0.67 for the shape branch, 0.64 for frequency anchoring and 0.52 for the sensor constraint, whereas the relative reductions of Table 3 are taken against different reference rows and are therefore not comparable with each other. Across these configurations the AUROC differences are comparable to the across-fold standard deviations, whereas the FPR95 reductions are unambiguous: the full method removes 42% of the FPR95 incurred by unconditional retrieval. Sensor filtering is inactive on single-channel targets such as JNU.

5 CONCLUSION

Pretrain once, deploy many times connects reusable vibration representations to training-free deployment: one frozen encoder serves unseen datasets and disjoint engines, while FAMB instantiates normality locally through Mel-frequency retrieval and bounded latent frequency-shape enhancement. With 100 samples from one normal engine, FAMB reaches 96.49%/96.88% AUROC/AUPR, exceeding the strongest baseline by 3.80/4.28 points. These results support a shift from repeatedly retraining machine-specific representations toward globally reusable ones with lightweight local normality modelling.

AI USE STATEMENT

In this work, we used generative AI tools to assist with translation of the manuscript from Chinese into English. We have not used generative AI tools to generate synthetic data sets, develop theoretical models or conceptual frameworks, formulate mathematical claims, assist in the writing of proofs, propose or refine hypotheses, design or provide feedback on the research methodology or experiments, implement methods, clean or reformat data sets, or interpret results; the remaining required disclosure tasks are not applicable to this work. Additionally, we used generative AI tools to edit the manuscript for readability. We have reviewed all AI-assisted work: every translated and edited passage was checked against the original manuscript by the authors, and all numerical results, table entries and claims were verified against the experimental logs. We take responsibility for the final content of this work, including text, claims and artefacts produced with the aid of generative AI.

ETHICS STATEMENT

This work uses only publicly released vibration datasets and de-identified engine acceptance-test records; it involves no human subjects, no personal data and no sensitive personal information. The liquid rocket engine measurements are used solely for health-state recognition, and no information identifying a specific manufacturer, site or production batch is reported. All public datasets are used

in accordance with their original licences and terms of use, and the sources are listed in Appendix A. The authors have read and adhere to the ICLR Code of Ethics.

Reproducibility statement

The data boundaries, splits and evaluation protocol are described in Sections 2 and 4.1; the input transform, backbone and scoring functions are given in Sections 3.1 and 3.2 as Eqs. (1)–(6). Appendix A lists all 38 pretraining sources with their persistent identifiers and states the isolation between the pretraining corpus and the evaluation data, so the cross-source corpus can be reassembled from the public releases. Appendix B separates what is shared by all compared methods—the windowing, STFT and normalisation chain, backbone interface and evaluation protocol—from the pretraining schedule and random seed of our encoder, every scoring hyperparameter of FAMB, and the objective-specific settings of the baselines (Appendix B.4), together with the repetition counts behind each table; within each of these groups no configuration is tuned per dataset, device or reference budget. The four public evaluation datasets are cited in Section 4.1 and are publicly available. Appendices C and D report the per-dataset and per-budget results, with standard deviations, underlying Figures 3 and 4 and Table 2.

References

Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., LeCun, Y., and Ballas, N. Self-supervised learning from images with a joint-embedding predictive architecture. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15619–15629, 2023.

Bardes, A., Ponce, J., and LeCun, Y. VICReg: Variance-invariance-covariance regularization for self-supervised learning. In *International Conference on Learning Representations (ICLR)*, 2022.

Batzner, K., Heckler, L., and König, R. EfficientAD: Accurate visual anomaly detection at millisecond-level latencies. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 128–138, 2024.

Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., and Joulin, A. Unsupervised learning of visual features by contrasting cluster assignments. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 9912–9924, 2020.

Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., and Joulin, A. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9650–9660, 2021.

Chen, F., Zhang, Y., Qin, Z., Fan, L., Jiang, R., Liang, Y., Wen, Q., and Deng, S. Learning multi-pattern normalities in the frequency domain for efficient time series anomaly detection. In *IEEE 40th International Conference on Data Engineering (ICDE)*, pp. 747–760, 2024.

Chen, S., Liu, Z., Li, C., Zou, D., He, X., and Zhou, D. Multi-mode fault diagnosis datasets of three-phase asynchronous motor under variable working conditions. *Data in Brief*, 2026.

Chen, X., Xie, S., and He, K. An empirical study of training self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9640–9649, 2021.

Cheng, H., Luo, J., Zhang, X., Liu, H., and Wu, F. Towards training noise-robust anomaly detection via collaborative adversarial flows. *Measurement*, 242:115865, 2025.

Davis, J. and Goadrich, M. The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, pp. 233–240, 2006.

Defard, T., Setkov, A., Loesch, A., and Audigier, R. PaDiM: A patch distribution modeling framework for anomaly detection and localization. In *Pattern Recognition. ICPR International Workshops and Challenges*, pp. 475–489. Springer, 2021.

Deng, H. and Li, X. Anomaly detection via reverse distillation from one-class embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9737–9746, 2022.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations (ICLR)*, 2021.

Eldele, E., Ragab, M., Qing, X., Edward, Chen, Z., Wu, M., Li, X., and Lee, J. UniFault: A fault diagnosis foundation model from bearing data. *arXiv preprint arXiv:2504.01373*, 2025.

Fang, Q., Lv, W., and Su, Q. AF-CLIP: Zero-shot anomaly detection via anomaly-focused CLIP adaptation. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM)*, 2025.

Fang, Y., Xie, J., Zhao, Y., Chen, L., Gao, Y., and Zheng, K. Temporal-frequency masked autoencoders for time series anomaly detection. In *IEEE 40th International Conference on Data Engineering (ICDE)*, pp. 1228–1241, 2024.

Feng, Y., Liu, Z., Chen, J., Lv, H., Wang, J., and Yuan, J. Make the rocket intelligent at IoT edge: Stepwise GAN for anomaly detection of LRE with multisource fusion. *IEEE Internet of Things Journal*, 9(4):3135–3149, 2022.

Fu, Y., Shao, Z., Yu, C., Li, Y., Xu, Y., Cheng, X., and Wang, F. Zeus: Towards tuning-free foundation model for time series analysis. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, 2026.

Gao, B.-B. and Wang, C. One language-free foundation model is enough for universal vision anomaly detection. *arXiv preprint arXiv:2601.05552*, 2026.

Gao, B.-B., Zhou, Y., Yan, J., Cai, Y., Zhang, W., Wang, M., Liu, J., Liu, Y., Wang, L., and Wang, C. AdaptCLIP: Adapting CLIP for universal visual anomaly detection. *arXiv preprint arXiv:2505.09926*, 2025.

Grill, J.-B., Strub, F., Altché, F., Tallec, C., Richemond, P. H., Buchatskaya, E., Doersch, C., Pires, B. A., Guo, Z. D., Azar, M. G., Piot, B., Kavukcuoglu, K., Munos, R., and Valko, M. Bootstrap your own latent: A new approach to self-supervised learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 21271–21284, 2020.

Guo, J., Lu, S., Zhang, W., Chen, F., Li, H., and Liao, H. Dinomaly: The less is more philosophy in multi-class unsupervised anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20405–20415, 2025.

Han, P., Liu, Z., He, X., Ding, S. X., and Zhou, D. Multi-condition fault diagnosis of dynamic systems: A survey, insights, and prospects. *IEEE Transactions on Automation Science and Engineering*, 22:15728–15744, 2025.

He, H., Qiao, H., Chen, Y., Yi, K., and Pang, G. TimeRadar: A domain-rotatable foundation model for time series anomaly detection. *arXiv preprint arXiv:2602.19068*, 2026.

He, K., Chen, X., Xie, S., Li, Y., Dollár, P., and Girshick, R. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16000–16009, 2022.

Hermary, R., Hicsonmez, S., Pineau, D., Shabayek, A. E. R., and Aouada, D. ASTER: Latent pseudo-anomaly generation for unsupervised time-series anomaly detection. In *International Conference on Pattern Recognition (ICPR)*, 2026.

Jung, W., Yun, S. H., Lim, Y. S., Cheong, S., and Park, Y. H. Vibration and current dataset of three-phase permanent magnet synchronous motors with stator faults. *Data in Brief*, 47:108952, 2023.

Khan, U., Capogrosso, L., Biondani, F., Magno, M., Fummi, F., Setti, F., and Cristani, M. ChronosAD: Leveraging time series foundation models for accurate anomaly detection. In *24th IEEE International Conference on Industrial Informatics (INDIN)*, 2026.

Lai, Z., Yang, C., Lan, S., Wang, L., Shen, W., and Zhu, L. BearingFM: Towards a foundation model for bearing fault diagnosis by domain knowledge and contrastive learning. *International Journal of Production Economics*, 275:109319, 2024.

Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., and Nandi, A. K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138:106587, 2020.

Li, F., Chen, J., Liu, Z., Lv, H., Wang, J., Yuan, J., and Xiao, W. A soft-target difference scaling network via relational knowledge distillation for fault detection of liquid rocket engine under multi-source trouble-free samples. *Reliability Engineering & System Safety*, 228:108759, 2022.

Li, J., Huang, R., Chen, Z., He, G., Gryllias, K. C., and Li, W. Deep continual transfer learning with dynamic weight aggregation for fault diagnosis of industrial streaming data under varying working conditions. *Advanced Engineering Informatics*, 55:101883, 2023.

Li, K., Ping, X., Wang, H., Chen, P., and Cao, Y. Sequential fuzzy diagnosis method for motor roller bearing in variable operating conditions based on vibration analysis. *Sensors*, 13(6):8013–8041, 2013.

Li, Q., Chen, B., Chen, Q., Li, X., Qin, Z., and Chu, F. HSE: A plug-and-play module for unified fault diagnosis foundation models. *Information Fusion*, 123:103277, 2025.

Li, Y.-F., Wang, H., and Sun, M. ChatGPT-like large-scale foundation models for prognostics and health management: A survey and roadmaps. *Reliability Engineering & System Safety*, 243:109850, 2024.

Lian, Z., Li, S., Huang, Q., Huang, Z., Liu, H., Qiu, J., Yang, P., and Tao, L. UBMF: Uncertainty-aware Bayesian meta-learning framework for fault diagnosis with imbalanced industrial data. *Knowledge-Based Systems*, 2025.

Liang, P., Deng, C., Wu, J., Yang, Z., Zhu, J., and Zhang, Z. Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform. *Computers in Industry*, 113:103132, 2019.

Lin, Y., Yan, H., Tong, X., Chang, Y., Wang, H., Zhou, Z., Gao, S., Wang, Y., and Zhang, W. Commonality in few: Few-shot multimodal anomaly detection via hypergraph-enhanced memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.

Liu, J., Peng, D., Wang, H., Liu, C., Li, Y.-F., and Xie, M. AeroGPT: Leveraging large-scale audio model for aero-engine bearing fault diagnosis. *IEEE Transactions on Cybernetics*, 2026. doi: 10.1109/TCYB.2026.3668256.

Liu, Z., Feng, Y., Chen, J., Wang, J., and Zhang, Z. Intelligent anomaly detection of liquid rocket engine with multi-source data. *Journal of Rocket Propulsion*, 48(3):79–86, 2022. (in Chinese). doi: 10.3969/j.issn.1672-9374.2022.03.010.

Lv, H., Chen, J., Wang, J., Yuan, J., and Liu, Z. A supervised framework for recognition of liquid rocket engine health state under steady-state process without fault samples. *IEEE Transactions on Instrumentation and Measurement*, 70:3518610, 2021.

Mannone, G., Fischer, P., and Dazer, M. Learning the language of vibration: A self-supervised transformer foundation model for PHM. *PHM Society European Conference*, 9(1):1–15, 2026.

Nie, Y., Nguyen, N. H., Sinthong, P., and Kalagnanam, J. A time series is worth 64 words: Long-term forecasting with transformers. In *International Conference on Learning Representations (ICLR)*, 2023.

Ren, C., Jiang, B., Lu, N., Simani, S., and Gao, F. Meta-learning with distributional similarity preference for few-shot fault diagnosis under varying working conditions. *IEEE Transactions on Cybernetics*, 54(5):2746–2756, 2024.

Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., and Gehler, P. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14318–14328, 2022.

Ruff, L., Vandermeulen, R. A., Görnitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E., and Kloft, M. Deep one-class classification. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pp. 4393–4402, 2018.

Sehri, M., Varejão, I., Hua, Z., Bonella, V., Santos, A., de Assis Boldt, F., Dumond, P., and Varejão, F. M. Towards a universal vibration analysis dataset: A framework for transfer learning in predictive maintenance and structural health monitoring. *International Journal of Prognostics and Health Management*, 2025.

Shao, S., McAleer, S., Yan, R., and Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Transactions on Industrial Informatics*, 15(4):2446–2455, 2019.

Singh, M. T. Graph-based fault diagnosis for rotating machinery: Adaptive segmentation and structural feature integration. *Results in Engineering*, 106566, 2025.

Wang, H., Liu, Z., Ge, Y., and Peng, D. Self-supervised signal representation learning for machinery fault diagnosis under limited annotation data. *Knowledge-Based Systems*, 239:107978, 2022.

Wang, J., Lv, H., Chen, J., Liu, Z., and Yuan, J. Intelligent detection method of liquid rocket engine health state. *Journal of Rocket Propulsion*, 47(4):52–58, 2021. (in Chinese). doi: 10.3969/j.issn.1672-9374.2021.04.008.

Wang, Y., Yu, Y., Sun, K., Lei, P., Zhang, Y., Zio, E., Xia, A., and Li, Y. RmGPT: A foundation model with generative pretrained transformer for fault diagnosis and prognosis in rotating machinery. *IEEE Internet of Things Journal*, 12(20):41562–41573, 2025.

Wen, G.-H. and Chen, K.-Y. Do time-series foundation models pay off for industrial monitoring? A cost-aware empirical study. *arXiv preprint arXiv:2608.22968*, 2026.

Wenig, P., Schmidl, S., and Papenbrock, T. Anomaly detectors for multivariate time series: The proof of the pudding is in the eating. In *IEEE 40th International Conference on Data Engineering Workshops (ICDEW)*, pp. 96–101, 2024.

Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., and Long, M. TimesNet: Temporal 2D-variation modeling for general time series analysis. In *International Conference on Learning Representations (ICLR)*, 2023.

Xia, H., Liu, Y., Hu, Y., Liu, P., and Dong, Z. Into the ORBIT for time series: Training regimes for foundation models. *arXiv preprint arXiv:2608.13262*, 2026.

Xiao, Y., Shao, H., Yan, S., Wang, J., Peng, Y., and Liu, B. Domain generalization for rotating machinery fault diagnosis: A survey. *Advanced Engineering Informatics*, 64:103063, 2025.

Xie, G., Wang, J., Liu, J., Lyu, J., Liu, Y., Wang, C., Zheng, F., and Jin, Y. IM-IAD: Industrial image anomaly detection benchmark in manufacturing. *IEEE Transactions on Cybernetics*, 2024.

Xu, H., Wang, Y., Jian, S., Liao, Q., Wang, Y., and Pang, G. Calibrated one-class classification for unsupervised time series anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, 2024.

Xu, J., Wu, H., Wang, J., and Long, M. Anomaly transformer: Time series anomaly detection with association discrepancy. In *International Conference on Learning Representations (ICLR)*, 2022.

Yao, X., Luo, Y., Qian, Z., and Zhang, C. ADPretrain: Advancing industrial anomaly detection via anomaly representation pretraining. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Yu, X., Zhao, Z., Zhang, X., Zhang, Q., Liu, Y., Sun, C., and Chen, X. Deep-learning-based open set fault diagnosis by extreme value theory. *IEEE Transactions on Industrial Informatics*, 18(1):185–196, 2022.

Zhang, W., Li, C., Peng, G., Chen, Y., and Zhang, Z. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing*, 100:439–453, 2018.

Zhang, X., Li, G., Henderson, P., Aragon-Camarasa, G., and Long, Z. Is training necessary for anomaly detection? *arXiv preprint arXiv:2601.22763*, 2026.

Zhao, C., Zio, E., and Shen, W. Domain generalization for cross-domain fault diagnosis: An application-oriented perspective and a benchmark study. *Reliability Engineering & System Safety*, 245:109964, 2024.

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., and Gao, R. X. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115:213–237, 2019.

Zhao, Z., Li, T., Wu, J., Sun, C., Wang, S., Yan, R., and Chen, X. Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. *ISA Transactions*, 107:224–255, 2020.

A Composition of the Pretraining Corpus

The cross-source pretraining corpus consists of 38 public vibration data sources, covering bearings, gears and gearboxes, rotor dynamics, electric motors, and rotating machinery scenarios such as slewing bearings, wind turbine drivetrains and turbofan test rigs. The sources differ considerably in sampling rate, channel count, speed and load conditions, and we perform no cross-source alignment of amplitude or frequency axis. Only vibration channels (acceleration and velocity) of each source are retained; non-vibration channels such as motor current, temperature and acoustic emission are discarded at conversion time, so the corpus is a pure vibration corpus even where a source provides other modalities. Each retained channel in every pkl file is split into an independent one-dimensional signal, uniformly cut into 2048-point windows (non-overlapping by default) and converted into a (1, 128, 128) log-Mel representation before entering pretraining. The total number of windows and the global mean and standard deviation are written into dataset_index.json by the data-generation script, from which the corpus scale can be reproduced. Pretraining is self-supervised and uses no health or fault labels, so the fault patterns contained in the corpus do not constitute label leakage.

Table A1 lists all sources with their persistent identifiers. The full address of a Mendeley entry in the table is `https://data.mendeley.com/datasets/<identifier>`, a Zenodo entry is `https://zenodo.org/records/<identifier>`, and a Figshare entry is `https://figshare.com/articles/dataset/<identifier>`.

Table A1: The 38 public vibration sources that make up the cross-source pretraining corpus.

ID	Dataset and subject	Modality and main operating conditions	Source (persistent identifier)
A1	CWRU rolling bearing dataset	Drive-end / fan-end acceleration; inner-race, outer-race and rolling-element faults, multiple speed–load conditions	GitHub: hitwzc/Bearing-datasets
A2	Paderborn University bearing data center	Acceleration + current; artificial and real damage of 6203 bearings, 32 operating conditions	mb.uni-paderborn.de/kat-datacenter
A3	UNSW bearing full-life dataset	Horizontal / vertical acceleration; natural evolution of spall severity, 6–20 Hz shaft frequency	Mendeley: h4df4mgrfb/3
A4	Ball bearing vibration and temperature full-life dataset	Acceleration + temperature; constant speed 1772 RPM run to failure	Mendeley: 5hcdd3tdvb/6

ID	Dataset and subject	Modality and main operating conditions	Source (persistent identifier)
A5	Politecnico di Torino rolling bearing test rig	Triaxial acceleration; roller / inner-race indentation 0–450 μm, 0–500 Hz $\times$ 4 load levels, including a 230 h endurance run	Zenodo: 3559553
A6	CITEF–UPM spherical roller bearing database	Acceleration; rolling-element-induced faults of FAG 22205E1KC3, 200–500 RPM	Zenodo: 3898942
A7	Variable-speed rolling bearing vibration and current (Subset 1)	Vibration + motor current; 680–2460 RPM	Mendeley: vxkj334rzv/7
A8	Variable-speed rolling bearing vibration and current (Subset 2)	Vibration + motor current; 680–2460 RPM	Mendeley: x3vhp8t6hg/7
A9	Variable-speed rolling bearing vibration and current (Subset 3)	Vibration + motor current; 680–2460 RPM	Mendeley: j8d8pfkvj2/7
A10	Time-varying-speed bearing vibration dataset	Acceleration; multiple health states $\times$ time-varying speed, 36 groups	Mendeley: v43hmbwxpm/1
A11	Ottawa UORED-VAFCLS bearing dataset	Acceleration + acoustics; inner race / outer race / ball / cage, three stages from healthy to developing to failed, 42 kHz	Mendeley: y2px5tg92h/5
A12	Cross-speed bearing vibration dataset	Acceleration; 5 health classes, 1000/1500/2000 r/min, 25.6 kHz	Zenodo: 20927070
A13	Compound-fault multi-domain bearing dataset (deep-groove ball bearing subset)	Acceleration; 31 single / compound faults $\times$ 6 speeds $\times$ 8/16 kHz	Mendeley: 53vtnjy6c6/1
A14	1205 EKTN9 bearing vibration data	Acceleration; bearing vibration under constant speed	Mostwiedzy: 221015219219252-0
A15	VBL-VA001 laboratory vibration dataset	Triaxial acceleration; normal / bearing fault / misalignment / unbalance, 20 kHz	Zenodo: 7006575
A16	5 MW wind turbine drivetrain bearing damage dataset	Simulated acceleration; floating spar, floating TLP and onshore platforms $\times$ healthy and two bearing damage classes $\times$ 6 environmental conditions	Zenodo: 10644756
A17	Wind turbine pitch bearing vibration and acoustic dataset	Acceleration + acoustic emission; time-varying load conditions	Mendeley: 2n9jkd384k/1
A18	UNSW TMCM gear crack dataset	Acceleration; gear cracks on a single-stage spur gear test rig	Mendeley: p72x3m92cv/1
A19	Aalto gear failure dataset	Acceleration; azimuth thruster gear failures, 16 kHz	Mendeley: 6cc4ctxthy/1
A20	DMTD electric vehicle reducer gear dataset	Dynamic meshing vibration; industrial measurements for manufacturing quality inspection	SciDB: 3544168ffcf7 4c4789e648ed070e21b2
A21	NREL wind turbine gearbox condition monitoring benchmark dataset	Acceleration; benchmark for wind turbine gearbox vibration analysis	OpenEI: submissions/738
A22	DGEN 380 turbofan engine test rig dataset	Triaxial acceleration + acoustics; includes a planetary gearbox, idle / medium / full-throttle conditions	Zenodo: 13860307
A23	Rotor–shaft test rig misalignment dataset	Acceleration; misalignment under varying load	Mendeley: cbpr9fvvwb/1
A24	Shaft angular misalignment dataset	Acceleration; horizontal / vertical angular misalignment, several severity levels	Mendeley: kf96jx9dzf/1
A25	Dual-rotor misalignment / looseness / rub-impact dataset	Acceleration; 10 operating conditions of a dual-rotor system	Mendeley: b6bfd7v86d/2
A26	Flexible rotor vibration dataset	Acceleration; time- and frequency-domain signals of a large flexible rotor simulation model	Mendeley: ny42v5cvt6/1
A27	Wind turbine rotor unbalance dataset	Output vibration only; 65 laboratory experiments	Dataverse: doi:10.34810/data2304

ID	Dataset and subject	Modality and main operating conditions	Source (persistent identifier)
A28	General rotating machinery fault vibration dataset	Acceleration + velocity; unbalance / misalignment / bearing fault × four speeds from 10 to 40 Hz, 6 kHz	Zenodo: 14388093
A29	Three-phase asynchronous motor multi-condition fault dataset	Acceleration; motor faults under variable speed and variable load	Mendeley: 6s3dggj9mw/1
A30	Three-phase PMSM stator fault dataset	Vibration at 25.6 kHz + three-phase current; turn-to-turn and coil-to-coil short circuits, several power levels	Mendeley: rgn5brrgrn/5
A31	PMSM multi-sensor vibration dataset	Triaxial acceleration; 5 fault classes × 4 sensor positions, including physical disturbances, 5 kHz	Zenodo: 19717375
A32	Smartphone-acquired induction motor fault dataset	Acceleration; multiple fault types × multiple speed–load conditions	Mendeley: rs4vz8n3t5/1
A33	Bearing and gearbox dataset for few-shot diagnosis	Acceleration; few-shot diagnosis scenarios for bearings and gearboxes	Mendeley: 2znkfwhzs7/1
A34	JUST slewing bearing dataset	Vibration + acoustic emission; 9 operating conditions of a slewing bearing	Mendeley: hwg8v5j8t6/1
A35	HVAC test rig belt looseness dataset	Three-phase current + vibration; looseness faults of a belt-driven system	Mendeley: 5tt3zb2pns/1
A36	SpectraQuest Machinery Fault Simulator (MFS) dataset	Acceleration; standard test-rig data with multiple channels and multiple fault types	www02.smt.ufrj.br/˜offshore/mfs
A37	Single and compound fault scenarios of rotating machinery	Acceleration; single-fault and double-fault scenarios	Figshare: 22693120
A38	BJTU-RAO subway train bogie dataset	Multi-sensor vibration; 1 health state and 50 faulty states under 9 operating conditions, 64 kHz	DOI: 10.1109/ACCESS.2025.3551603

Isolation statement. The four public evaluation sources of Section 4 (JNU, XJTU-G, SEU-G, HUST) and the 16 liquid rocket engines are not part of the corpus above; pretraining and evaluation do not overlap at the level of datasets or of raw records. After pretraining, the backbone is frozen, and each target scene builds its memory and statistics only from its own normal references.

B Hyperparameter Settings

The input representation, the windowing and the evaluation protocol of Section B.1 are shared by every method reported in the paper, including all baselines. The remaining settings are separated by role: Sections B.2 and B.3 specify our own pipeline—the cross-source encoder and the FAMB scorer—while the objective-specific settings of the self-supervised baselines and of the trained detection heads are listed in Section B.4. Within each of these groups, no hyperparameter is tuned for any target dataset, target device or reference budget, and no anomalous sample or additional calibration set takes part in parameter selection. The values below are exactly those of our implementation.

B.1 Input representation and backbone

Item	Value
Input size	128×128 log-Mel spectrogram
Window length	2048 samples, non-overlapping windowing by default
STFT	$N_{\mathrm{FFT}} = 512$, HOP = 16, 128 Mel filters (Slaney scale)

Item	Value
Time-axis handling	STFT with centre padding (of the signal, inside the transform) gives 1 + 2048 / 16 = 129 frames; the last frame is cropped to give exactly 128 columns. The spectrogram time axis itself is not padded: all 128 columns are computed frames, and no zero or constant column is appended
Normalisation chain	channel-level z-score → per-window z-score → PowerToDB (reference = window maximum)
Backbone	ViT: patch 16×16, hidden 256, 8 layers, 8 heads, FFN 1024, no CLS token
Parameter count	about 6.4 M
Patch grid	8×8 (frequency rows H = 8, time columns W = 8; each row covers 16 Mel bins)
Feature normalisation	parameter-free LayerNorm (plain_ln), comparable across layers
Global descriptor	mean pooling of the layer-7 patch tokens followed by L2 normalisation

B.2 Cross-source pretraining

Item	Value
Pretraining objective	DINO-style multi-view self-distillation, projection head output dimension 4096
View composition	2 global views + 4 local views (64×64)
Teacher update	EMA, momentum m = 0.996
Teacher temperature	warmup from 0.04 to a target of 0.07 over the first 3 epochs of the 20-epoch schedule
Epochs / batch size	20 epochs, batch size 128
Optimiser learning rate	initial 2.5×10^{-4}, cosine decay to 1×10^{-6}, warmup 10 epochs
Weight decay / gradient clipping	cosine schedule 0.04 → 0.4; gradient clipping 3.0
Random seed	42
Export	the teacher backbone is exported at the end of training and frozen in all downstream experiments

B.3 FAMB detection and scoring

Item	Value
Frequency neighbourhood radius ρ_f	1 (the frequency band of the query plus its two immediate neighbours)
Enhancement weight λ	0.5
Normal margin γ	0.0 (the normal mean is the onset of curvature anomaly)
Temperature τ	1.5
Global pre-filter top_k	8 (all reference windows when fewer than 8 are available)
Stabilisation term ε	1×10^{-12}
Shape-branch replicate unit	(reference window, time column); W = 8 columns per window at 128×128 with patch size 16
Normal-scale estimator	leave-one-replicate nearest-curvature distances; at n = 1 this reduces to leave-one-column, giving W distances
Shrinkage pseudo-count κ	8.0
Pooled σ floor	5% of the pooled mean, guarding against $\sigma_2 \to 0$
Time-column aggregation	mean over the W columns
Shape branch weights	(point, slope, curve) = (0, 0, 1); only the second-order curvature branch is kept
Participating layers	layers 1, 3, 5 and 7; each layer builds its own memory, fused with equal weights
Three-axis constraints	sensor strictly identical; local frequency neighbourhood; time matched freely

Item	Value
Channel mode	separate (multi-channel records encoded channel by channel, sensor identity retained)
Patch aggregation	the window-level score is the mean over all patches
Boundary frequency rows	no second-order difference is computed, the residual is 0 and the enhancement coefficient is 0
Memory precision	float16

Evaluation-side settings. Figure 3 and Tables C1–C2 use 5 independent repetitions, with Gaussian noise at SNR = 5.0 dB added to test windows. Table 2 and Table C3 use 20 repetitions with fresh normal-reference draws and no added noise. Figure 4, Table 3 and Table D1 use 3-fold device-level cross-validation without added noise. Seeds and repetition counts are fixed in the implementation. Reference draws and noise realisations follow the corresponding protocol; the test set is fixed across repeated public evaluations and within each engine fold. These evaluation-side settings apply identically to our method and to every baseline in Tables C1–C3 and Tables 3–D1.

B.4 Baseline-specific settings

The baselines share with FAMB everything specified in Section B.1—input size, windowing, STFT and normalisation chain, backbone architecture and interface—and the split definitions, normal-reference budgets, test sets and repetition counts of Section 4.1 and the evaluation-side paragraph above. Only the objective-specific choices differ, and none of them is tuned per target dataset, target device or reference budget.

Self-supervised baselines (DINO, MAE, I-JEPA, MoCo v3, BYOL, SwAV and VICReg, Table C1) are pretrained from scratch on the normal-training split of the single target dataset they are evaluated on, using the objective, projection head and loss weights of their reference formulation under the shared backbone capacity of Section B.1, with the same optimiser family, learning-rate schedule and random seed as Section B.2, trained for 100 epochs until the self-supervised loss converged. Capacity is therefore matched and every baseline is trained to convergence, and the baselines additionally keep the target-specific advantage noted in Section 4.2: each one is pretrained on the very dataset it is evaluated on, whereas our encoder is pretrained once on Appendix A and never observes any of them.

Detection heads (Table C2) operate on the frozen features of one shared backbone. RD4AD-style, Dinomaly-style and Deep SVDD-style heads are the trained ones: they are optimised on the target normal references only, with no access to anomalous samples or test windows. PaDiM-style needs no gradient update—it fits position-dependent Gaussian statistics on the same references—and Global kNN, PatchCore-style and RAD-style, like FAMB, perform no target-domain gradient update either. Head-specific choices follow the reference formulation of each method and are kept fixed across all four target datasets, all reference budgets and all engine folds, so every entry in Tables C1–D1 differs only in the mechanism being compared.

C Full Per-dataset Results

Tables C1 and C2 give the per-dataset values summarised in Figure 3, and Table C3 gives the per-dataset values summarised in Table 2. All implementations share the same input representation, backbone interface, reference set, test set and score aggregation.

Table C1: Large-scale cross-source pretraining against conventional single-source pretraining, with the FAMB head fixed (per-dataset form of Figure 3a). Each backbone occupies two rows; each dataset cell is the mean ± standard deviation of 5 independent runs, and the macro-average column is the equally weighted mean over the four datasets. Test windows carry Gaussian noise at SNR = 5.0 dB. None of the four evaluation sources took part in our cross-source pretraining; the single-source baselines were pretrained separately on the normal-training split of the corresponding target dataset. The row in bold is our method.

Backbone / pretraining	**Metric**	**JNU**	**XJTU-G**	**SEU-G**	**HUST**	**Macro-avg**
DINO	AUROC	97.64 ± 0.36	98.92 ± 0.05	97.88 ± 0.19	97.94 ± 0.41	98.10
	AUPR	97.81 ± 0.31	98.91 ± 0.05	97.96 ± 0.16	97.97 ± 0.42	98.16
MAE	AUROC	97.32 ± 0.47	98.90 ± 0.05	97.64 ± 0.24	97.90 ± 0.34	97.94
	AUPR	97.57 ± 0.39	98.90 ± 0.05	97.77 ± 0.22	97.96 ± 0.33	98.05
I-JEPA	AUROC	97.50 ± 0.36	98.90 ± 0.06	98.12 ± 0.13	98.02 ± 0.39	98.14
	AUPR	97.72 ± 0.30	98.90 ± 0.06	98.17 ± 0.12	98.05 ± 0.39	98.21
MoCo v3	AUROC	97.29 ± 0.53	98.94 ± 0.02	98.70 ± 0.12	97.30 ± 0.39	98.06
	AUPR	97.52 ± 0.44	98.94 ± 0.02	98.72 ± 0.12	97.51 ± 0.31	98.17
SwAV	AUROC	96.41 ± 0.40	96.63 ± 0.11	96.83 ± 0.17	96.62 ± 0.38	96.62
	AUPR	96.60 ± 0.34	96.65 ± 0.12	96.79 ± 0.22	96.73 ± 0.27	96.69
VICReg	AUROC	98.14 ± 0.23	98.99 ± 0.01	98.92 ± 0.03	97.56 ± 0.45	98.40
	AUPR	98.05 ± 0.28	98.99 ± 0.01	98.91 ± 0.05	98.26 ± 0.29	98.55
BYOL	AUROC	97.54 ± 0.32	98.88 ± 0.03	98.99 ± 0.01	98.64 ± 0.22	98.51
	AUPR	97.72 ± 0.29	98.89 ± 0.03	98.99 ± 0.01	98.65 ± 0.22	98.56
Ours	**AUROC**	**99.41 ± 0.21**	**99.48 ± 0.02**	**100.00 ± 0.00**	**99.71 ± 0.16**	**99.65**
	AUPR	**99.39 ± 0.25**	**99.52 ± 0.02**	**100.00 ± 0.00**	**99.72 ± 0.14**	**99.66**

Table C2: Detection heads with our pretrained backbone fixed (per-dataset form of Figure 3b). Heads that require training are allowed to optimise on the target normal references; FAMB, PatchCore-style, RAD-style, Global kNN and PaDiM-style perform no target-domain gradient update. Test windows carry Gaussian noise at SNR = 5.0 dB. The random-weight row does not fix the backbone and is listed for reference only.

Detection head	**Metric**	**JNU**	**XJTU-G**	**SEU-G**	**HUST**	**Macro-avg**
Random weights + FAMB (backbone not loaded)	AUROC	52.78 ± 3.20	86.88 ± 0.50	96.82 ± 0.57	65.13 ± 1.57	75.40
	AUPR	54.64 ± 3.07	87.25 ± 0.47	96.91 ± 0.48	66.58 ± 1.42	76.35
Global kNN	AUROC	97.87 ± 0.29	98.53 ± 0.28	99.66 ± 0.24	97.98 ± 0.77	98.51
	AUPR	98.05 ± 0.33	98.76 ± 0.32	99.54 ± 0.35	98.27 ± 0.65	98.66
PatchCore-style	AUROC	98.91 ± 0.50	98.63 ± 0.29	100.00 ± 0.00	99.53 ± 0.07	99.27
	AUPR	98.93 ± 0.46	98.78 ± 0.25	100.00 ± 0.00	99.59 ± 0.06	99.33
PaDiM-style	AUROC	98.17 ± 0.66	97.38 ± 0.38	99.35 ± 0.20	95.92 ± 1.26	97.71
	AUPR	98.18 ± 0.69	97.79 ± 0.19	99.22 ± 0.29	96.50 ± 1.15	97.92
RAD-style	AUROC	97.66 ± 0.21	97.17 ± 0.64	100.00 ± 0.00	99.60 ± 0.26	98.61
	AUPR	97.92 ± 0.29	97.43 ± 0.49	100.00 ± 0.00	99.62 ± 0.23	98.74
Deep SVDD-style	AUROC	86.32 ± 0.90	88.02 ± 1.92	99.56 ± 0.12	88.66 ± 4.39	90.64
	AUPR	88.47 ± 0.69	91.58 ± 0.67	99.54 ± 0.14	91.91 ± 3.03	92.88
RD4AD-style	AUROC	97.87 ± 0.44	97.75 ± 0.74	99.81 ± 0.16	97.85 ± 0.76	98.32
	AUPR	98.05 ± 0.45	98.15 ± 0.42	99.79 ± 0.19	98.34 ± 0.60	98.58
Dinomaly-style	AUROC	99.09 ± 0.43	98.77 ± 0.54	99.97 ± 0.01	97.96 ± 0.55	98.95
	AUPR	99.15 ± 0.41	98.96 ± 0.39	99.97 ± 0.01	98.36 ± 0.33	99.11
FAMB	**AUROC**	**99.41 ± 0.21**	**99.48 ± 0.02**	**100.00 ± 0.00**	**99.71 ± 0.16**	**99.65**
	AUPR	**99.39 ± 0.25**	**99.52 ± 0.02**	**100.00 ± 0.00**	**99.72 ± 0.14**	**99.66**

Table C3: Few-shot results for $n = 1$, 2 and 4 normal references (per-dataset form of Table 2). Each dataset cell is the mean $\pm$ standard deviation of 20 independent repetitions with the references redrawn each time and no added noise. The test set is fixed, and the few-shot references are drawn from a complete normal reference pool covering all operating conditions.

Head / n	**Metric**	**JNU**	**XJTU-G**	**SEU-G**	**HUST**	**Macro-avg**
PatchCore-style / 1	AUROC	96.05 ± 3.26	91.22 ± 7.87	93.23 ± 2.90	89.30 ± 1.84	92.45
	AUPR	96.76 ± 2.34	93.20 ± 5.50	93.84 ± 2.35	91.80 ± 1.37	93.90
RAD-style / 1	AUROC	97.69 ± 1.23	93.80 ± 4.56	95.60 ± 1.98	88.78 ± 1.25	93.97
	AUPR	97.85 ± 1.20	94.74 ± 3.39	95.88 ± 1.74	91.14 ± 1.17	94.90
Dinomaly-style / 1	AUROC	96.13 ± 1.46	90.95 ± 6.81	95.06 ± 1.82	85.79 ± 4.14	91.98
	AUPR	96.63 ± 1.26	93.76 ± 4.68	95.33 ± 1.74	89.14 ± 3.28	93.72
FAMB / 1	**AUROC**	**98.25 ± 0.75**	**96.21 ± 3.77**	**98.22 ± 0.69**	**92.30 ± 1.25**	**96.25**
	AUPR	**98.35 ± 0.73**	**96.81 ± 2.80**	**98.32 ± 0.61**	**94.20 ± 0.91**	**96.92**
PatchCore-style / 2	AUROC	96.69 ± 2.36	96.28 ± 3.28	99.72 ± 0.20	91.47 ± 1.50	96.04
	AUPR	97.05 ± 2.34	96.87 ± 2.71	99.70 ± 0.22	93.37 ± 1.17	96.75
RAD-style / 2	AUROC	97.98 ± 0.79	95.68 ± 3.26	99.69 ± 0.26	89.83 ± 0.73	95.80
	AUPR	98.26 ± 0.68	96.30 ± 2.72	99.68 ± 0.26	91.97 ± 0.66	96.55
Dinomaly-style / 2	AUROC	97.71 ± 0.44	92.94 ± 2.02	99.66 ± 0.30	88.51 ± 1.15	94.70
	AUPR	97.99 ± 0.42	95.23 ± 1.35	99.64 ± 0.30	91.18 ± 0.93	96.01
FAMB / 2	**AUROC**	**98.36 ± 0.60**	**98.71 ± 1.28**	**99.78 ± 0.25**	**93.13 ± 0.94**	**97.50**
	AUPR	**98.56 ± 0.54**	**98.78 ± 1.21**	**99.78 ± 0.24**	**94.80 ± 0.77**	**97.98**
PatchCore-style / 4	AUROC	98.60 ± 0.65	98.42 ± 0.87	99.84 ± 0.17	92.01 ± 1.22	97.22
	AUPR	98.67 ± 0.62	98.58 ± 0.74	99.83 ± 0.18	93.20 ± 0.97	97.57
RAD-style / 4	AUROC	99.17 ± 0.48	98.01 ± 1.10	99.82 ± 0.20	91.31 ± 1.08	97.08
	AUPR	99.16 ± 0.50	98.16 ± 0.94	99.80 ± 0.21	92.46 ± 0.86	97.40
Dinomaly-style / 4	AUROC	98.19 ± 0.81	92.57 ± 1.59	99.78 ± 0.22	89.11 ± 2.75	94.91
	AUPR	98.28 ± 0.77	95.09 ± 1.04	99.76 ± 0.24	90.79 ± 2.33	95.98
FAMB / 4	**AUROC**	**99.49 ± 0.35**	**99.61 ± 0.24**	**99.89 ± 0.12**	**94.62 ± 1.29**	**98.40**
	AUPR	**99.49 ± 0.35**	**99.61 ± 0.24**	**99.87 ± 0.14**	**95.64 ± 0.95**	**98.65**

D Additional Figures and Tables

Table D1: Cross-device results under 3-fold device-level cross-validation for different budgets of normal reference engines, without added noise. Each cell is the 3-fold mean ± standard deviation; bold marks the best result in the column. Figure 4 plots the same results.

Detection head	Metric	8 engines / 800 samples	4 engines / 400 samples	2 engines / 200 samples	1 engine / 100 samples
PatchCore-style	AUROC	99.26 ± 0.42	98.62 ± 0.77	96.69 ± 1.56	92.69 ± 2.38
	AUPR	99.38 ± 0.38	98.54 ± 0.64	96.06 ± 1.12	92.60 ± 2.57
RAD-style	AUROC	99.02 ± 0.67	98.84 ± 0.96	97.22 ± 1.34	88.62 ± 2.85
	AUPR	99.04 ± 0.66	98.61 ± 0.95	97.16 ± 1.39	88.11 ± 2.78
RD4AD-style	AUROC	97.19 ± 1.16	96.17 ± 1.42	89.49 ± 2.78	85.42 ± 3.99
	AUPR	97.53 ± 1.11	96.63 ± 1.34	89.74 ± 2.47	85.53 ± 3.84
Dinomaly-style	AUROC	96.93 ± 1.01	92.83 ± 2.21	87.07 ± 3.84	69.27 ± 5.72
	AUPR	97.05 ± 0.93	92.12 ± 2.48	87.58 ± 3.12	70.47 ± 5.12
FAMB	**AUROC**	**99.59 ± 0.54**	**98.96 ± 0.71**	**97.57 ± 0.88**	**96.49 ± 1.98**
	AUPR	**99.51 ± 0.56**	**99.04 ± 0.74**	**98.04 ± 0.92**	**96.88 ± 2.14**

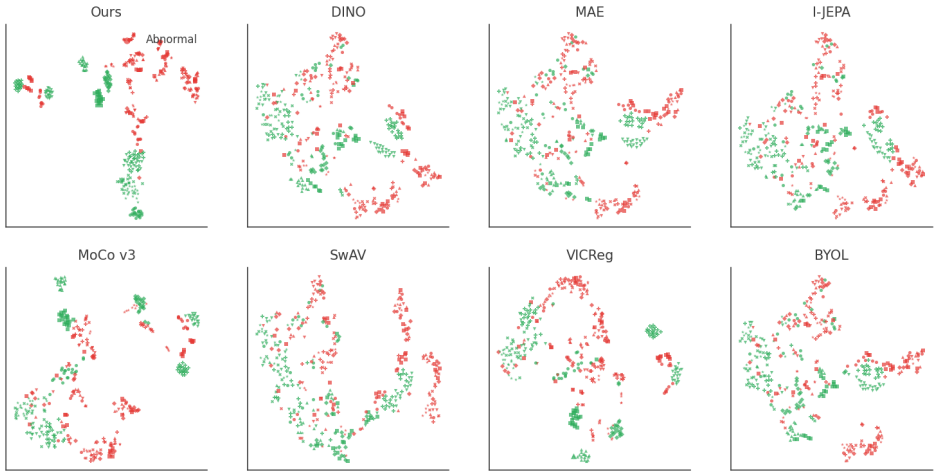


Figure 5: t-SNE projections of frozen window-level features of the HUST test windows (colour = health state), one panel per encoder. Our backbone (top left) is pretrained once on the 38-source corpus of Appendix A and never observes HUST; the seven baseline panels are the self-supervised baselines of Section 4.2—DINO, MAE, I-JEPA, MoCo v3, SwAV, VICReg and BYOL—each pretrained from scratch on the HUST normal-training split (400 normal samples) under the protocol of Section 4.2. Our features separate the two health states more clearly, whereas the baseline representations show greater overlap between them. The projection is computed from frozen features for visualisation only: no memory, statistic or threshold is derived from it, and Figure 3 evaluates the same encoders quantitatively.

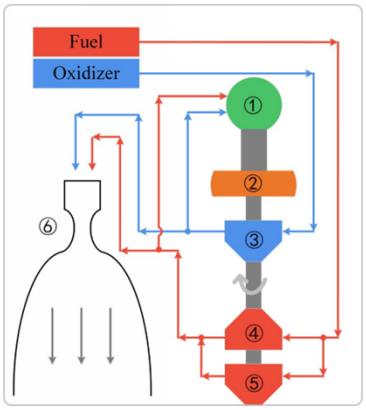


Figure 6: Main structure and working principle of a liquid rocket engine (① gas generator, ② main turbine, ③ oxidiser pump, ④ first-stage fuel pump, ⑤ second-stage fuel pump, ⑥ thrust chamber). The 11 vibration sensors of Section 4.4 cover ①, ③, ⑤, ⑥ and the oxidiser pre-pump, which is not marked separately; ② and ④ carry no sensors.