Supplementary Material

TRAC-AD: Transformation-Response-Aware Correction for Anomaly Detection

Yunli Chen, Yinghua Zhu, and Rui Tian

S1 Seeds and Implementation

Table S1 lists the memory and paired-response seeds s_m and s_r for configurations $c = 0, 1, 2$, with offset $o_c = 1000003c$.

Table S1: Random seeds.

c	Offset	Memory	Paired response
0	0	17	3407
1	1000003	1000020	1003410
2	2000006	2000023	2003413

For fitting-image index i, source-tile record index j, transformation index t, and branch index b (S1: 0; native: 1), use $s_m + 1009j$ for memory rows, $s_r + h(a) + 1009i + 101b$ for response positions, and $s_r + h(a) + 1000i + t$ for transformed views. Here $h(a) = 0$ for AD 2 categories and $h(a) = 100003 \sum_{v \in a} \mathrm{ord}(v)$ otherwise, using the source identifiers AeBAD-S, PCB, and MetalParts. The transformation order is brightness, gamma, contrast, RGB gain, spatial gradient, vignette, and soft shadow; transformed-view seeds are shared across scales.

L4 parallel-analysis seeds are $20260827 + o_c + 10009u + 101b$ for the three sources above, indexed by u in that order, and $20260823 + o_c + 10009v + 101b$ for AD 2 categories indexed by v in the order can, fabric, fruit jelly, rice, sheet metal, vial, wallplugs, walnuts. All indices start at zero. Image index i follows the fixed fitting-image list for each source. Tile index j runs cumulatively across that list, with tiles ordered from top to bottom and left to right within each image.

S1 resizes the image short side to 448 pixels. The native branch uses 896 pixels for memory construction and the original image resolution for queries and paired-response extraction. Memory-row indices are shared by the fused and four layer banks. After processing N valid tokens, each branch retains $\mathrm{round}(rN)$ memory rows, with $r = 0.1$ for S1 and $r = 1/44$ for the native branch. Response positions are sampled from concatenated valid tile-token occurrences, retaining overlapping tile occurrences.

S2 Per-Configuration Results

Tables S2 and S3 report the per-configuration metrics for main Tables 1(a) and 5, respectively. Base and TRAC-AD results are reported separately for the main evaluation and the residual-scoring controls, with methods paired within each experiment and configuration.

Table S2: Per-configuration local PRO@0.05 for main Table 1(a). Scores are percentages.

Setting	c	Base	TRAC-AD	Gain (pp)
AeBAD-S	0	68.81204	72.17337	+3.36133
AeBAD-S	1	69.67342	72.68897	+3.01556
AeBAD-S	2	69.08547	72.29771	+3.21224
RobustAD-derived	0	80.62750	81.81510	+1.18761
RobustAD-derived	1	80.73098	81.86759	+1.13661
RobustAD-derived	2	80.50406	81.58921	+1.08515
TESTpub	0	58.18586	58.46206	+0.27620
TESTpub	1	58.69134	59.05783	+0.36649
TESTpub	2	57.75503	58.57780	+0.82277

Table S3: Per-configuration scores for main Table 5.

Setting / protocol	c	Base	L4-Scalar	Normal-PCA	TRAC-AD
AeBAD-S illumination: local	0	0.7368272463454307	0.731791687510492	0.7559661487016668	0.7612722057467302
AeBAD-S illumination: local	1	0.7418815616105021	0.7348977583258829	0.7582672127488231	0.7636464146876947
AeBAD-S illumination: local	2	0.7381144979584421	0.7319112573498295	0.7566372502557324	0.7616065456737862
TESTpub: local	0	0.5818480953972911	0.5683700868856784	0.5826414439253493	0.5846146407260568
TESTpub: local	1	0.5869074510337172	0.5737116682673973	0.5890008778679178	0.5905770018103801
TESTpub: local	2	0.5775507273010152	0.5662037679775862	0.5847438627207854	0.5857720752676276
TESTpub: original resolution	0	0.6313418004321739	0.6220884030655532	0.634160877862126	0.6380696608371963
TESTpub: original resolution	1	0.6323433280276287	0.6253737008629222	0.6401323494209588	0.6415892162047243
TESTpub: original resolution	2	0.6307264561989407	0.6238491799713248	0.6394997446597159	0.6416496905014215

Local: PRO@0.05. Original resolution: AucPro@0.05. Values are proportions on a 0–1 scale.

S3 Calibration Statistics

Tables S4 and S5 summarize source calibration for the main and control experiments. Entries are means over configurations 0–2; inference uses each configuration's own calibration statistics and thresholds.

Table S4: Source-calibration medians and 99.5th percentiles.

		Main Table 1		Main Table 5			
Source	Scale	Base	TRAC-AD	Base	L4-Scalar	Normal-PCA	TRAC-AD
AeBAD-S	S1	0.5797 / 0.8608	0.5275 / 0.8284	0.5797 / 0.8608	0.5334 / 0.7983	0.5257 / 0.8127	0.5275 / 0.8284
AeBAD-S	Native	0.7646 / 0.9713	0.6900 / 0.9157	0.7646 / 0.9713	0.6950 / 0.8843	0.6865 / 0.9028	0.6900 / 0.9157
PCB	S1	0.3865 / 0.6435	0.3630 / 0.6245	—	—	—	—
PCB	Native	0.7762 / 0.9765	0.6922 / 0.9489	—	—	—	—
MetalParts	S1	0.5403 / 0.7410	0.5055 / 0.6930	—	—	—	—
MetalParts	Native	0.6965 / 0.8774	0.6315 / 0.8178	—	—	—	—
can	S1	0.4262 / 0.8242	0.4128 / 0.8067	0.4262 / 0.8242	0.4122 / 0.7929	0.4065 / 0.7857	0.4128 / 0.8067
can	Native	0.6502 / 0.9013	0.6213 / 0.8727	0.6502 / 0.9013	0.6204 / 0.8585	0.6114 / 0.8513	0.6213 / 0.8727
fabric	S1	0.7071 / 0.8423	0.6532 / 0.7788	0.7071 / 0.8423	0.6529 / 0.7762	0.6463 / 0.7672	0.6532 / 0.7788
fabric	Native	0.8166 / 0.9358	0.7596 / 0.8778	0.8166 / 0.9358	0.7608 / 0.8659	0.7518 / 0.8618	0.7596 / 0.8778
fruit jelly	S1	0.2497 / 0.8139	0.2288 / 0.7876	0.2497 / 0.8139	0.2350 / 0.7755	0.2303 / 0.7691	0.2288 / 0.7876
fruit jelly	Native	0.4352 / 0.9109	0.3719 / 0.8822	0.4352 / 0.9109	0.3888 / 0.8554	0.3722 / 0.8624	0.3719 / 0.8822
rice	S1	0.7192 / 0.8136	0.6704 / 0.7538	0.7192 / 0.8136	0.6707 / 0.7496	0.6625 / 0.7417	0.6704 / 0.7538
rice	Native	0.7824 / 0.8905	0.7173 / 0.8210	0.7824 / 0.8905	0.7172 / 0.8116	0.7120 / 0.8126	0.7173 / 0.8210
sheet metal	S1	0.7593 / 0.8872	0.7002 / 0.8200	0.7593 / 0.8872	0.7000 / 0.8153	0.6915 / 0.8089	0.7002 / 0.8200
sheet metal	Native	0.7554 / 0.9139	0.6957 / 0.8441	0.7554 / 0.9139	0.6954 / 0.8396	0.6886 / 0.8349	0.6957 / 0.8441
vial	S1	0.2277 / 0.8384	0.2120 / 0.8252	0.2277 / 0.8384	0.2198 / 0.8120	0.2149 / 0.8019	0.2120 / 0.8252
vial	Native	0.6000 / 0.9483	0.5291 / 0.9277	0.6000 / 0.9483	0.5375 / 0.8738	0.5234 / 0.9006	0.5291 / 0.9277
wallplugs	S1	0.5870 / 0.8081	0.5525 / 0.7646	0.5870 / 0.8081	0.5503 / 0.7586	0.5426 / 0.7470	0.5525 / 0.7646
wallplugs	Native	0.7224 / 0.9168	0.6619 / 0.8671	0.7224 / 0.9168	0.6659 / 0.8471	0.6608 / 0.8465	0.6619 / 0.8671
walnuts	S1	0.4961 / 0.8356	0.4515 / 0.8082	0.4961 / 0.8356	0.4609 / 0.7863	0.4492 / 0.7791	0.4515 / 0.8082
walnuts	Native	0.7389 / 0.9283	0.6734 / 0.8812	0.7389 / 0.9283	0.6832 / 0.8603	0.6752 / 0.8542	0.6734 / 0.8812

Each entry pairs the means of m_s and q_s over configurations 0–2, in raw map-score units. A dash indicates an unevaluated combination.

Table S5: Source-derived fused-score thresholds.

	Main Table 1		Main Table 5			
Source	Base	TRAC-AD	Base	L4-Scalar	Normal-PCA	TRAC-AD
AeBAD-S	0.8969 / 1.4923	0.9270 / 1.4893	0.8969 / 1.4923	0.9029 / 1.4853	0.9300 / 1.5052	0.9270 / 1.4893
PCB	0.8100 / 1.5371	0.8510 / 1.4677	—	—	—	—
MetalParts	0.8564 / 1.5121	0.8507 / 1.5107	—	—	—	—
can	0.9536 / 1.3313	0.9594 / 1.3237	0.9535 / 1.3313	0.9568 / 1.3298	0.9606 / 1.3467	0.9594 / 1.3236
fabric	0.8279 / 1.3956	0.8465 / 1.4255	0.8279 / 1.3956	0.8393 / 1.3934	0.8382 / 1.4404	0.8465 / 1.4255
fruit jelly	0.9649 / 1.2397	0.9685 / 1.2213	0.9649 / 1.2397	0.9664 / 1.2406	0.9676 / 1.2494	0.9685 / 1.2213
rice	0.7726 / 1.8055	0.7898 / 1.9563	0.7726 / 1.8055	0.7837 / 1.8032	0.7896 / 2.0132	0.7898 / 1.9563
sheet metal	0.8990 / 1.4036	0.9090 / 1.4204	0.8990 / 1.4036	0.9081 / 1.4006	0.9077 / 1.4430	0.9090 / 1.4204
vial	0.9612 / 1.2565	0.9648 / 1.2381	0.9612 / 1.2565	0.9644 / 1.2426	0.9647 / 1.2455	0.9648 / 1.2381
wallplugs	0.8815 / 1.4474	0.8943 / 1.4082	0.8815 / 1.4474	0.8866 / 1.4077	0.8933 / 1.4448	0.8943 / 1.4082
walnuts	0.9418 / 1.2588	0.9539 / 1.2428	0.9418 / 1.2588	0.9506 / 1.2362	0.9536 / 1.2672	0.9539 / 1.2428

Each entry pairs the means of θ_{pix} and θ_{img} over configurations 0–2, in fused-score units.

S4 Response-Basis Specifications

Table S6 gives the dimensions and ranks of all 66 L4 bases.

Table S6: L4 response-basis ranks. Every basis has shape $1024 \times k$.

Source	S1 k_0	S1 k_1	S1 k_2	Native k_0	Native k_1	Native k_2
AeBAD-S	128	127	128	145	145	144
PCB	75	78	77	109	108	106
MetalParts	102	103	102	135	136	136
can	71	71	71	109	108	109
fabric	114	113	112	113	113	113
fruit jelly	87	86	87	100	99	99
rice	89	89	87	108	104	105
sheet metal	138	138	140	146	154	150
vial	77	77	77	109	109	109
wallplugs	97	97	98	125	125	121
walnuts	102	102	102	128	131	130

S5 Additional Evaluation Protocols

The sensitivity analysis evaluates all 1,084 TESTpub images and all 273 AeBAD-S illumination images with configurations 0–2. Rank, transformation amplitude, and transformation family are varied independently. Each scoring variant is calibrated on the source-normal calibration split. Rank zero recovers Base, while the default uses the rank selected by parallel analysis. Amplitude factors scale deviations from the identity; global-only and spatial-only variants use four and three transformation types, respectively. The performance tables report means and sample standard deviations across configurations, with paired differences formed within each configuration before averaging.

The frozen perturbation experiment evaluates the complete illumination domain. Brightness, gamma, blur, and noise act on RGB images before preprocessing, with the source response bases, memory banks, calibration statistics, and thresholds held fixed. Noise realizations are shared between methods and configurations. The capacity experiment uses nested uniform subsets of aligned memory rows at 100%, 50%, and 25% capacity. All capacities retain the original response bases and fit their calibration statistics using the same 51 source-normal calibration images.

For CORAL-L2, a fixed label-agnostic split provides 55 unlabeled target images for covariance estimation and 218 disjoint images for evaluation. All original feature dimensions are retained. With covariance regularization $\lambda = 1$, the source-to-target transform is $(C_s + I)^{-1/2}(C_t + I)^{1/2}$ for row-vector features. It acts on the detector's L_2-normalized source reference features and source calibration descriptors; evaluation queries retain their original normalized target representation. Transformed source features are used directly for distance computation. Each fused or layer feature space has its own covariance transform. Source covariance is estimated from each original memory bank; target covariance uses all valid patch occurrences from the adaptation images, including tile overlaps. The adaptation split contains 44 anomalous and 11 normal images; all 55 images contribute without label filtering. Base and TRAC-AD use the same held-out evaluation set.

The controlled capacity timing experiment uses one RTX 3090 Ti, configuration 0, and three fixed decoded source RGB images. Three warmup inputs precede five repetitions of RGB-to-fused-score inference per input, giving 15 observations per method and capacity. This interval includes preprocessing, backbone extraction, two-branch retrieval, residual scoring, stitching, and scale fusion. A separate measurement uses pre-extracted features and ten repetitions per input. Methods alternate order, and CUDA operations are synchronized. Timing variability is calculated over repeated measurements; performance variability is calculated over the three source configurations. Decoding and disk I/O are outside both timing intervals.

S6 Response Spectra Across Sources and Scales

Table S7 lists the covariance traces, participation ratios, and selected ranks for all nine sources and both branches. These statistics use the eigenvalues from the original source-response parallel analysis, with 100 row-permutation replicates and a 95th-percentile null threshold for each eigenvalue. The same spectra define the normalized-eigenvalue curves in Figure S1 and the observed-to-null ratios in Figure S2. They characterize the centered L4 response distribution used for correction.

Table S7: L4 source-response spectra. Values are means ± sample SD over three source configurations. Trace measures centered response energy; the participation ratio measures spectral concentration.

Source	Branch	Covariance trace	Participation ratio	PA rank
can	S1	0.04351 ± 0.00038	19.79 ± 0.28	71.0 ± 0.0
can	native	0.03727 ± 0.00032	39.25 ± 0.68	108.7 ± 0.6
fabric	S1	0.01822 ± 0.00023	45.33 ± 0.55	113.0 ± 1.0
fabric	native	0.00707 ± 0.00003	44.17 ± 0.76	113.0 ± 0.0
fruit_jelly	S1	0.09424 ± 0.00236	31.34 ± 0.65	86.7 ± 0.6
fruit_jelly	native	0.16607 ± 0.00210	33.41 ± 0.54	99.3 ± 0.6
rice	S1	0.01860 ± 0.00022	27.74 ± 0.17	88.3 ± 1.2
rice	native	0.00449 ± 0.00014	42.27 ± 0.97	105.7 ± 2.1
sheet_metal	S1	0.00564 ± 0.00006	97.45 ± 2.25	138.7 ± 1.2
sheet_metal	native	0.00521 ± 0.00036	103.84 ± 16.32	150.0 ± 4.0
vial	S1	0.12682 ± 0.00089	26.05 ± 0.24	77.0 ± 0.0
vial	native	0.19322 ± 0.00036	44.27 ± 0.25	109.0 ± 0.0
wallplugs	S1	0.01091 ± 0.00033	43.78 ± 0.66	97.3 ± 0.6
wallplugs	native	0.00690 ± 0.00029	56.97 ± 3.63	123.7 ± 2.3
walnuts	S1	0.01053 ± 0.00013	43.68 ± 0.71	102.0 ± 0.0
walnuts	native	0.00790 ± 0.00010	66.51 ± 2.67	129.7 ± 1.5
AeBAD-S	S1	0.04084 ± 0.00043	70.95 ± 0.81	127.7 ± 0.6
AeBAD-S	native	0.04087 ± 0.00014	95.26 ± 0.85	144.7 ± 0.6

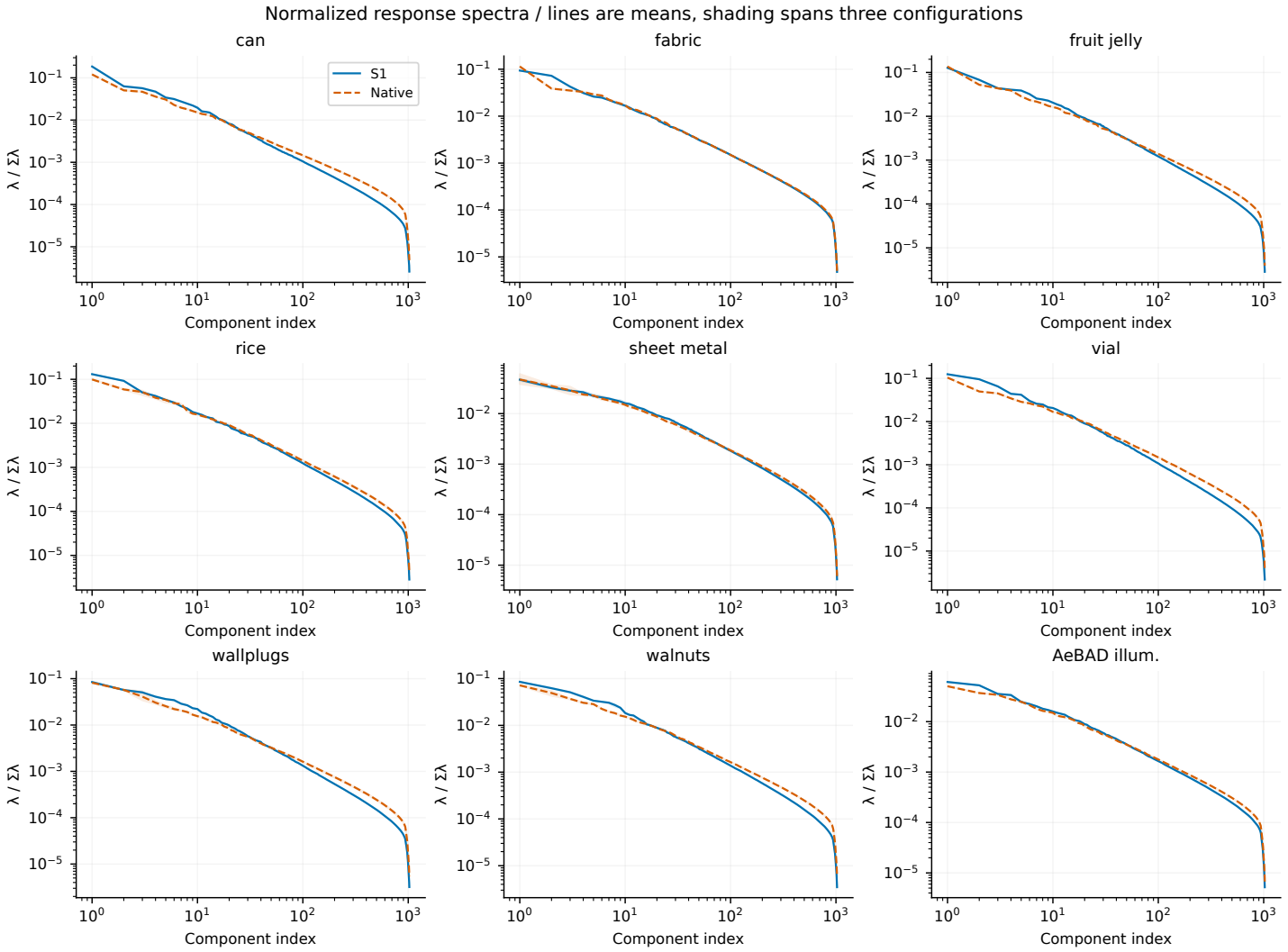

Figure S1: Normalized L4 response covariance spectra for the eight TESTpub categories and AeBAD-S, with S1 and native branches. Curves summarize the distribution of response energy across directions.

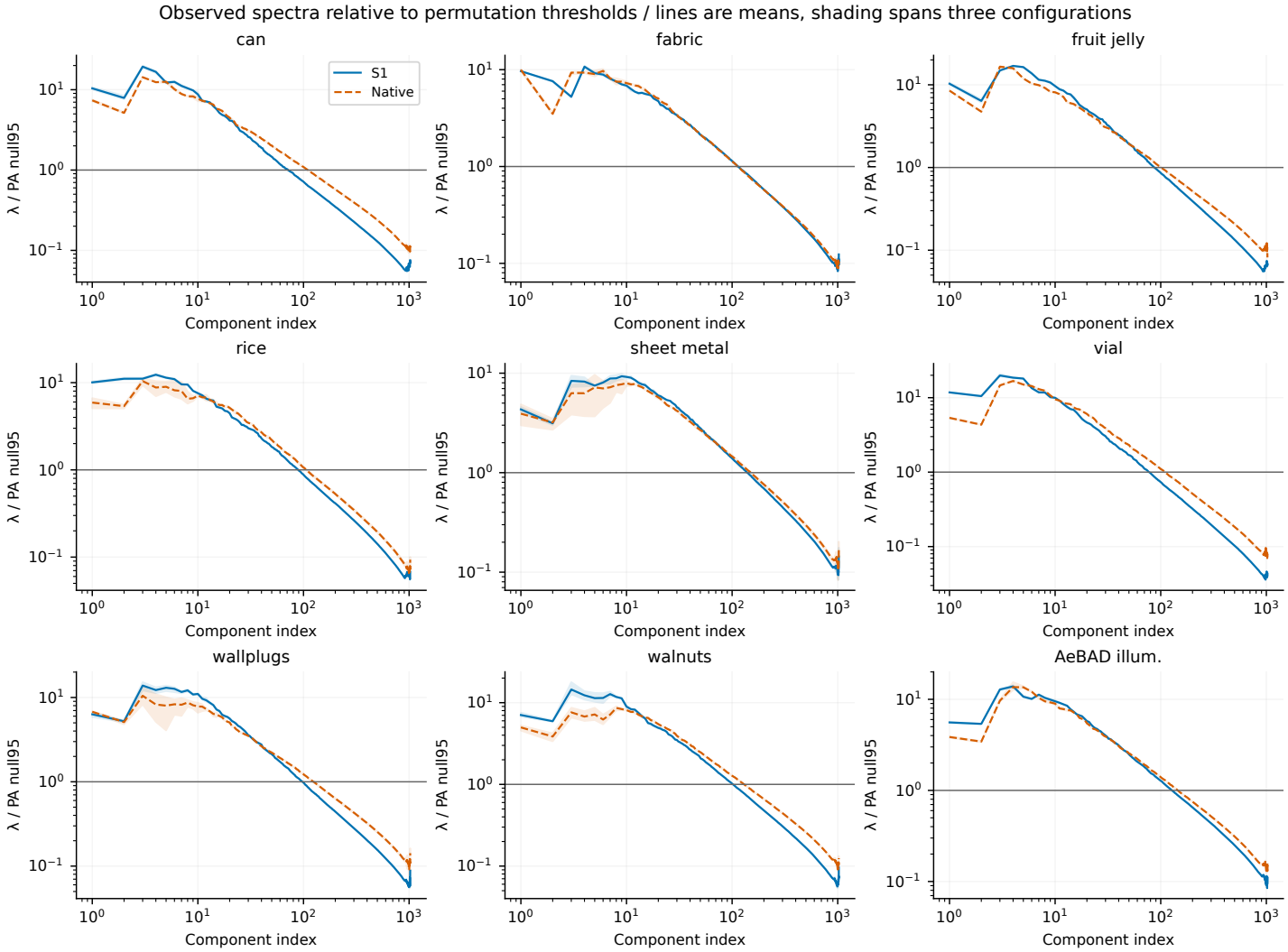

Figure S2: Observed response eigenvalues relative to their 95th-percentile permutation thresholds. The parallel-analysis rank follows the contiguous leading directions that exceed the threshold.

S7 Layer-Wise Regional Energy

Table S8 reports the projected-energy fraction for normal and defect regions on each layer. For a region, the fraction is the sum of squared projected residual norms divided by the sum of squared residual norms over valid tile-token occurrences. The normal/defect ratio compares these two fractions; the gap is their difference. Each layer uses its own source-selected rank. Defect tokens intersect the resized positive mask, and overlapping tile occurrences remain distinct.

Table S8: Regional projected-energy fractions on AeBAD-S illumination, using the final fused-space reference for every layer. Normal and defect fractions are percentages. Ratios and gaps are paired within configuration before averaging. Values are means ± sample SD over three configurations.

Branch	Layer	Normal (%)	Defect (%)	Normal/defect ratio	Gap (pp)
S1	L1	28.53 ± 1.66	21.97 ± 1.05	1.30 ± 0.02	6.56 ± 0.71
S1	L2	53.86 ± 0.10	35.34 ± 0.18	1.52 ± 0.01	18.52 ± 0.18
S1	L3	52.12 ± 0.07	34.18 ± 0.29	1.53 ± 0.01	17.94 ± 0.27
S1	L4	53.14 ± 0.21	34.73 ± 0.33	1.53 ± 0.02	18.41 ± 0.50
native	L1	21.67 ± 0.43	17.88 ± 0.50	1.21 ± 0.01	3.79 ± 0.06
native	L2	52.64 ± 0.14	43.15 ± 0.15	1.22 ± 0.00	9.49 ± 0.10
native	L3	56.46 ± 0.03	43.28 ± 0.04	1.30 ± 0.00	13.18 ± 0.02
native	L4	60.87 ± 0.10	45.47 ± 0.23	1.34 ± 0.00	15.40 ± 0.14

S8 Image-Level Errors and Case Selection

The image-level analysis includes all 426 TESTpub images from can, wallplugs, and sheet_metal and all three source configurations. For each of the 1,278 image–configuration pairs, we compute TP, FP, and FN counts at each method's source-calibrated threshold. Normal-image pixel FPR pools pixels from defect-free images; defect recall pools positive ground-truth pixels from anomalous images. Precision uses all images. The local evaluation grid has a longest side of at most 224 pixels. Pixel AP and post-hoc maximum F1 are computed from the full category score distribution; the latter is used to inspect the attainable precision–recall balance.

For each category, we select one normal image with decreased pixel FPR, one with increased pixel FPR, one anomalous image with decreased recall, and one with increased recall. Cases are selected near the median three-configuration mean change within each directional pool. For the normal-image cases and the anomalous wallplugs and sheet_metal cases, eligible images have the same nonzero direction of change in all three configurations. For the two anomalous can cases, eligibility is determined by the direction of the three-configuration mean. For each selected image, we display the configuration whose change is closest to this mean among configurations with the same direction; configuration index breaks ties.

Figures S3–S5 present the 12 selected cases. RGB images show the input appearance; masks and score maps use the local evaluation grid. For each method, scores are divided by its source-calibrated threshold and displayed on a common scale from 0 to 2, with values above 2 saturated. White contours mark the decision threshold and cyan contours mark ground truth. The accompanying data files provide all three configurations, selected image identifiers, and per-image counts.

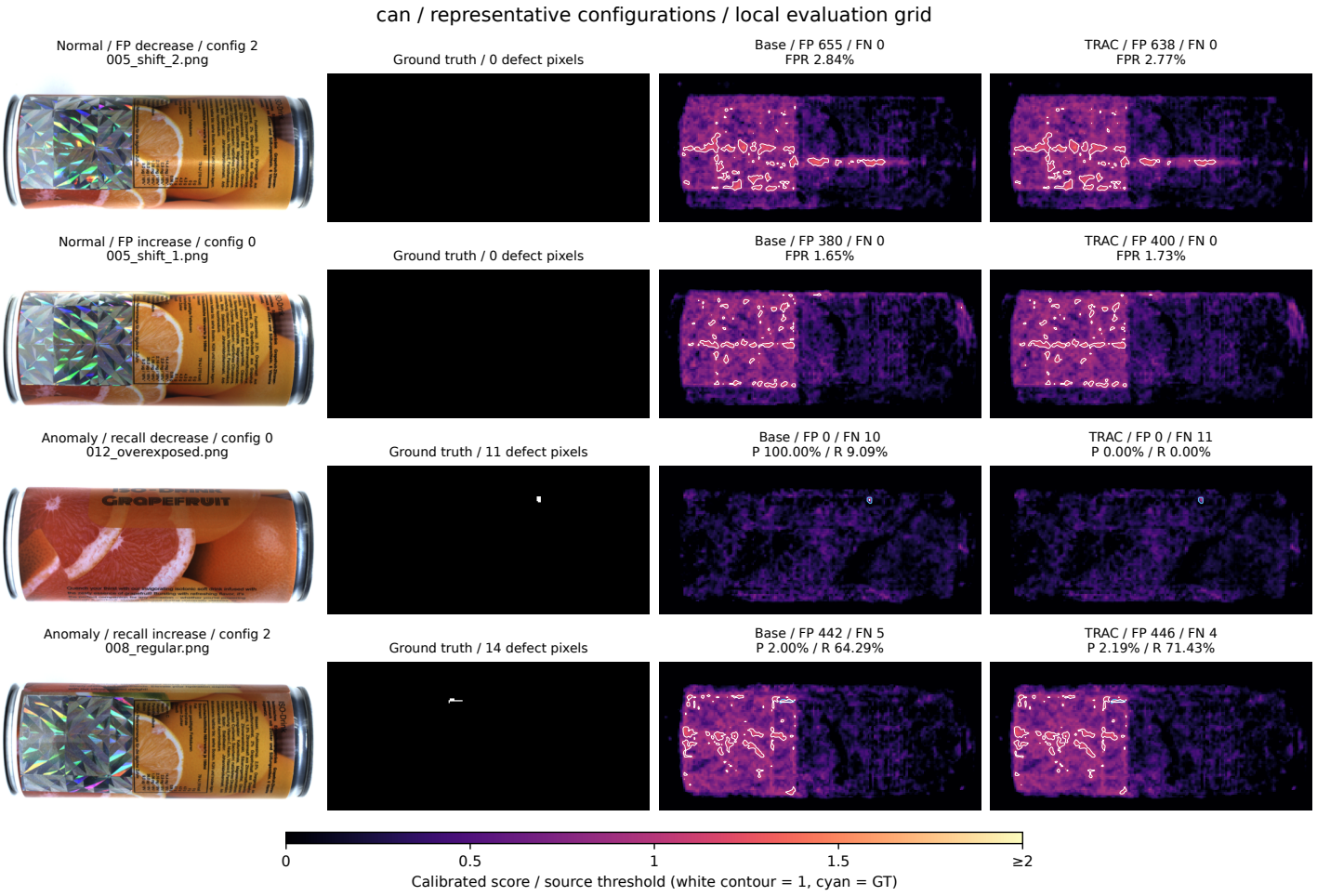

Figure S3: Can examples selected by the stated median-change procedure. The four rows show normal-image FP decrease, normal-image FP increase, anomaly recall decrease, and anomaly recall increase.

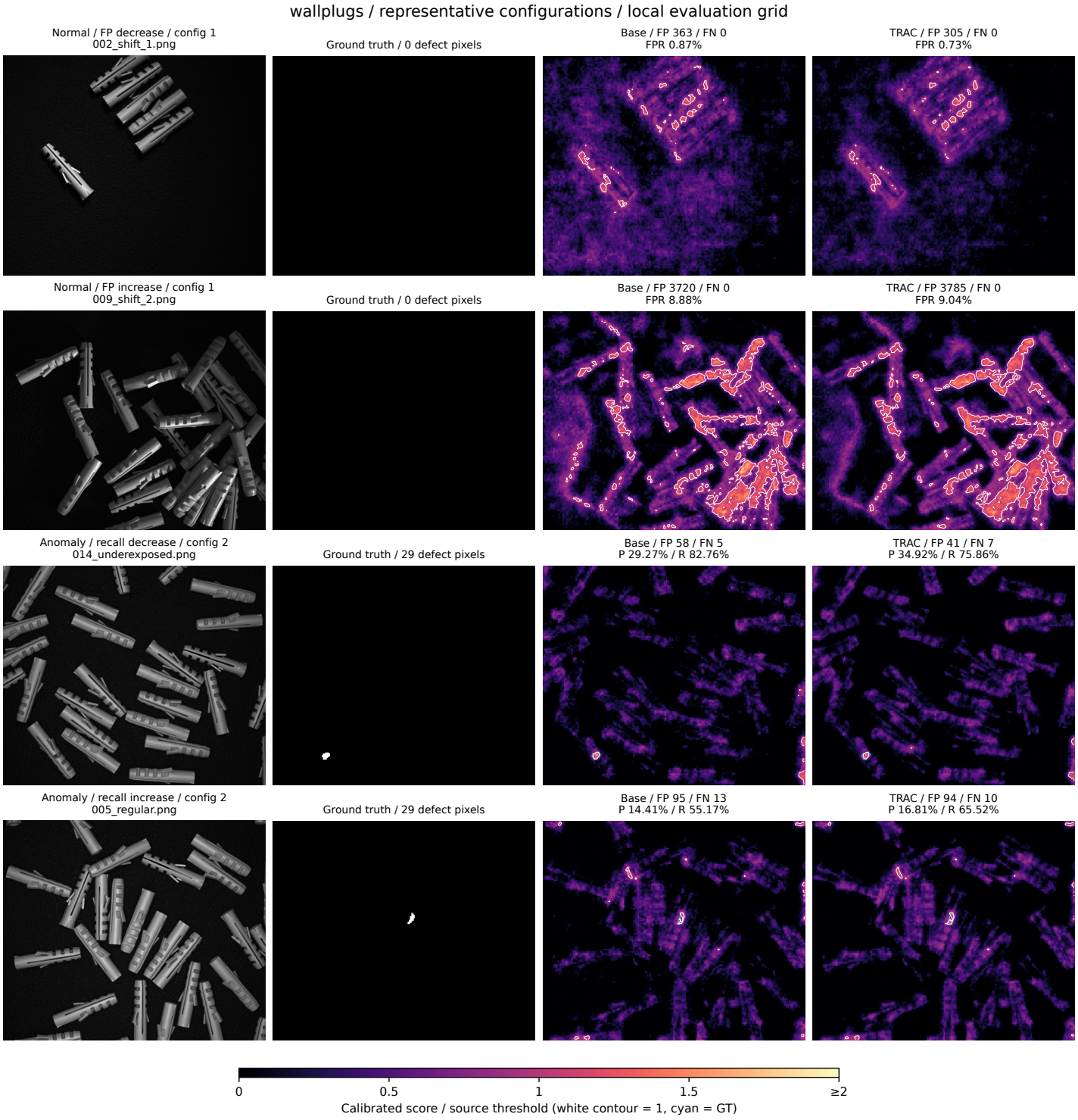

Figure S4: Wallplugs examples selected by the same procedure, with shared score-to-threshold scaling. The four rows show normal-image FP decrease, normal-image FP increase, anomaly recall decrease, and anomaly recall increase.

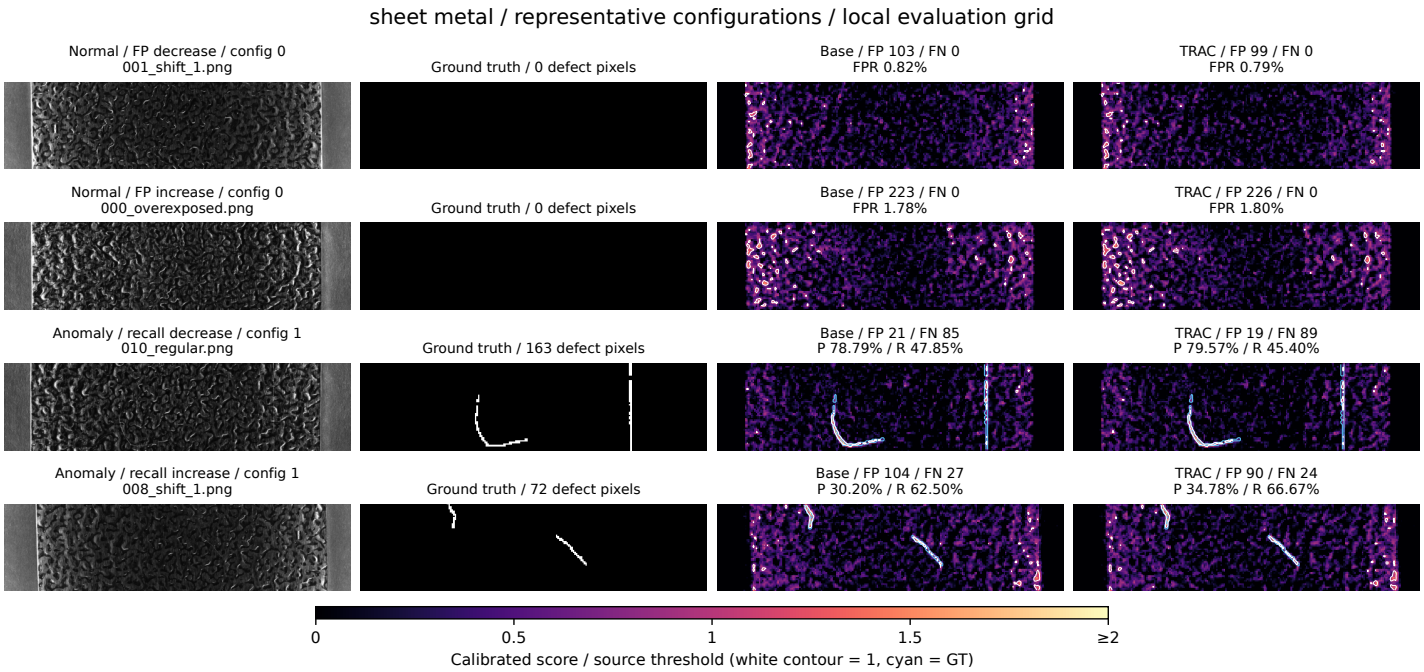

Figure S5: Sheet_metal examples selected by the same procedure, with shared score-to-threshold scaling. The four rows show normal-image FP decrease, normal-image FP increase, anomaly recall decrease, and anomaly recall increase.