Spectral-Spatial Extraction through Layered Tensor Decomposition for Hyperspectral Anomaly Detection

Minru Bai

Hunan University

Joint work with Quan Yu, Yu-Hong Dai

第二届数学与人工智能科学大会

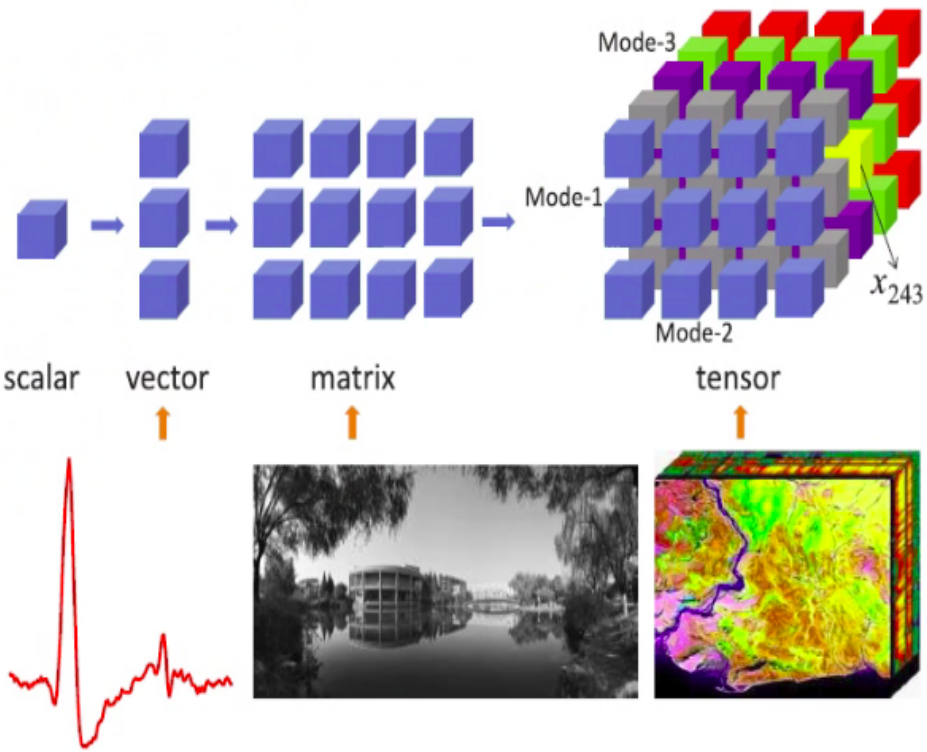
Contents

Contents

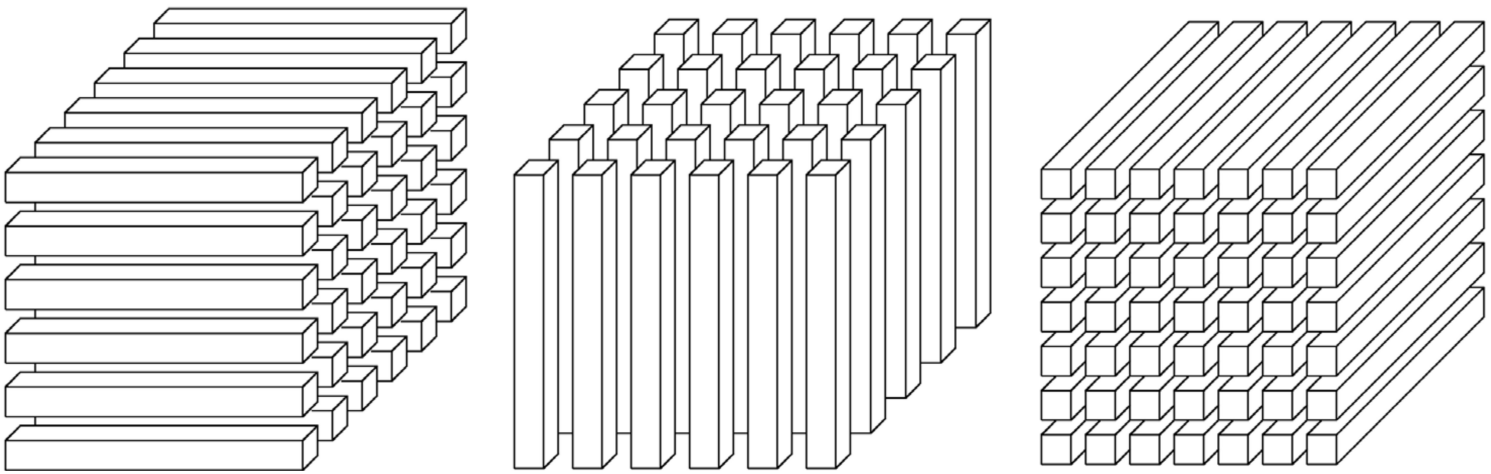
Tensor

There are a large number of high-dimensional data:

- Video data;
- Hyperspectral data;
- Traffic data;
- ...



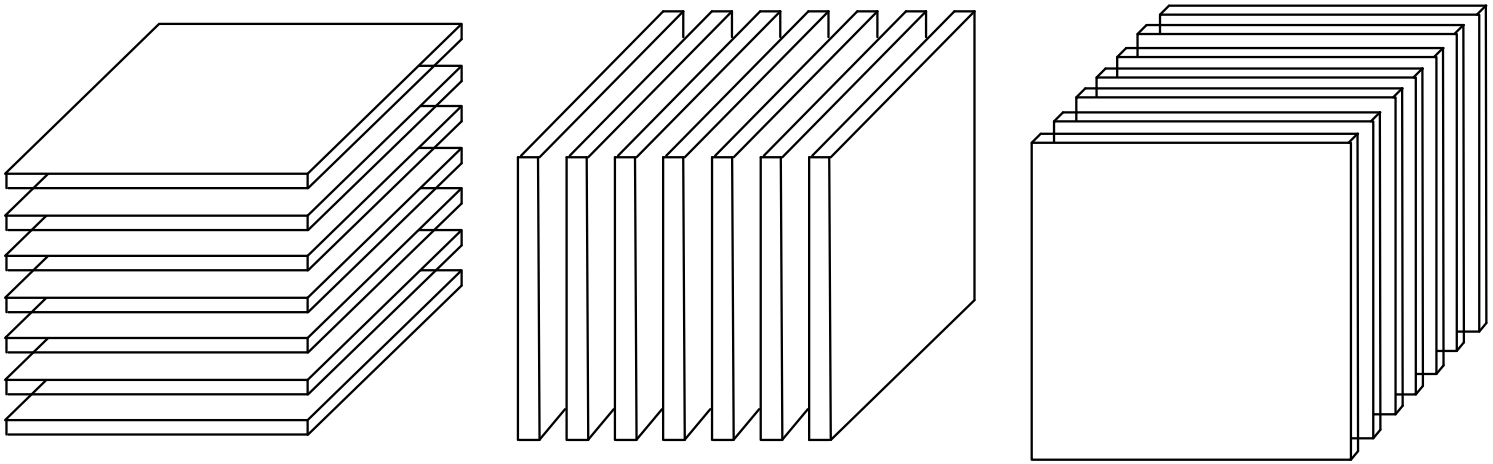
Fibers



(a) row fibers $\mathcal{X}(i,:,k)$ (b) column fibers $\mathcal{X}(:,j,k)$ (c) tube fibers $\mathcal{X}(i,j,:)$

Figure 1: Fibers of a 3rd-order tensor.

Slice



(a) Horizontal slices $\mathcal{X}(i,:,:)$

(b) Lateral slices $\mathcal{X}(:,j,:)$

(c) Frontal slices $\mathcal{X}(:,:,k)$

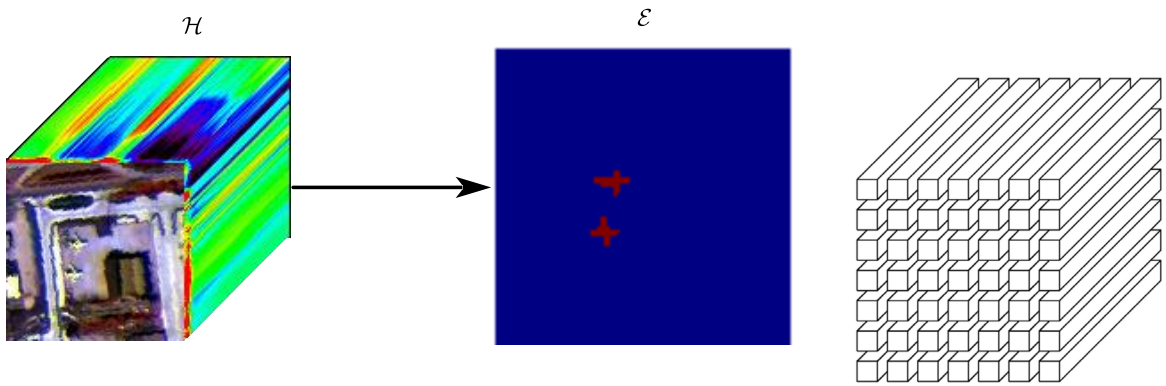
Figure 2: Slices of a 3rd-order tensor.

LRTR model

Given an observed Hyperspectral image (HSI) $\mathcal{H}$, the low rank tensor representation (LRTR) model for hyperspectral anomaly detection (HAD) can be expressed as

$$\text{[Spatial Anomaly]} \min_{\mathcal{Z},\mathcal{E}} \operatorname{rank}(\mathcal{Z}) + \lambda \|\mathcal{E}\|_{\ell_{F,0}}{}^{1}, \quad \text{s.t.} \quad \mathcal{H} = \mathcal{D} * \mathcal{Z} + \mathcal{E}.$$

$$\text{[Spectral Anomaly]} \min_{\mathcal{Z},\mathcal{E}} \operatorname{rank}(\mathcal{Z}) + \lambda \|\mathcal{E}\|_{\ell_{F,0}}, \quad \text{s.t.} \quad \mathcal{H} = \mathcal{Z} \times_3 D + \mathcal{E}.$$



[1] $\|\mathcal{E}\|_{\ell_{F,0}} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \|\mathcal{S}(i,j,:)\|_F^0$

- Spatial Anomaly
  - TLRA-MSL[2]: ε-TNN + $\ell_{F,1}$
  - PCA-TLRSR[3]: Weighted-TNN + $\ell_{F,1}$
- Spectral Anomaly
  - LARTVAD [4]: TNN + $\ell_{F,1}$ + $\|\mathcal{Z}\|_{TV}$

① The spectral domain anomaly or spatial domain anomaly is often neglected;
② Large-scale matrix singular value decomposition (SVD) is involved.

[2]He, Xu, et al. "Anomaly Detection for Hyperspectral Imagery via Tensor Low-Rank Approximation With Multiple Subspace Learning." IEEE Transactions on Geoscience and Remote Sensing (2023).

[3]Wang, Minghua, et al. "Learning tensor low-rank representation for hyperspectral anomaly detection." IEEE Transactions on Cybernetics 53.1 (2022): 679-691.

[4]Sun, Liu, et al. "Hyperspectral anomaly detection with tensor average rank and piecewise smoothness constraints." IEEE TNNLS 34.11 (2022): 8679-8692.

Contents

To address the aforementioned limitations, we propose an HAD model based on layered tensor decomposition (LTD), which incorporates two synergistically integrated layers. LTD efficiently ② detects both spectral and spatial anomaly ① within a unified framework, as illustrated in Figure 3.

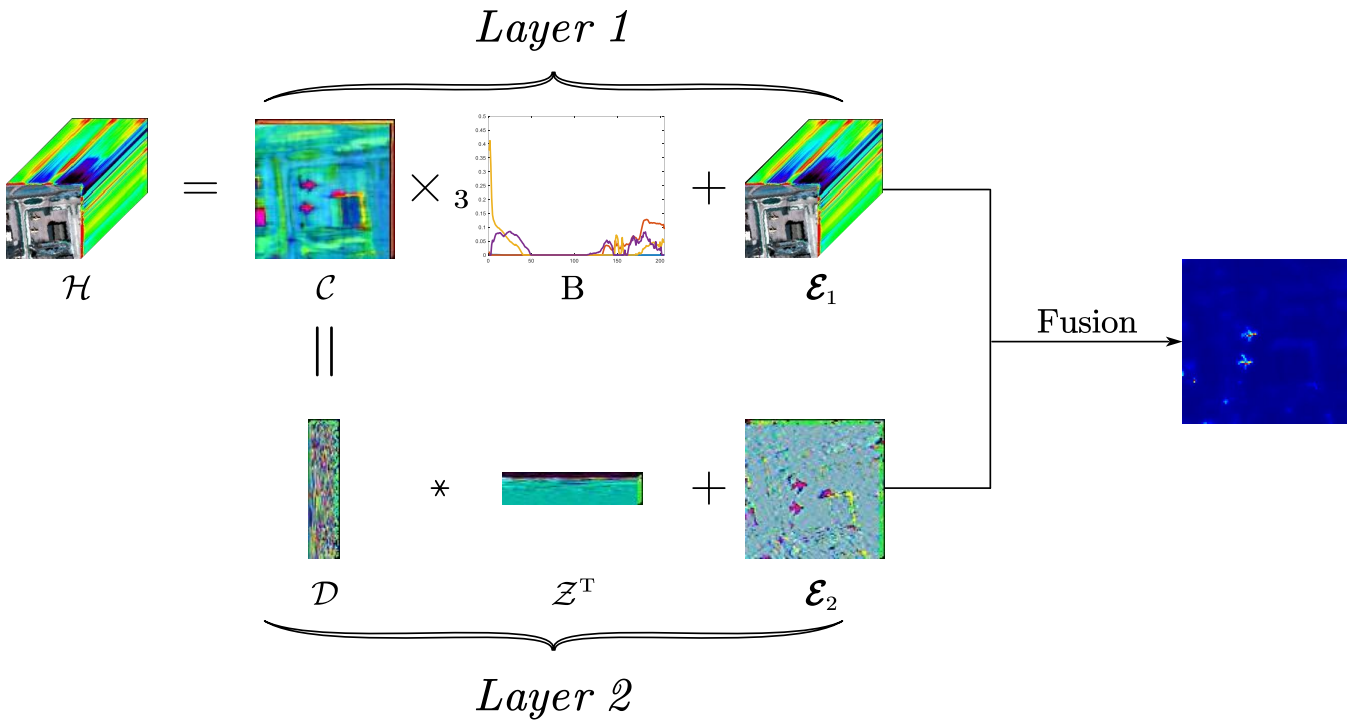


Figure 3: Flowchart of the proposed anomaly detection method.

Layer 1: Spectral anomaly extraction

Hyperspectral images (HSIs) have significant data redundancy due to high correlation between their spectral bands. Exploiting this, the first layer of our model decomposes the HSI into a low-dimensional background and a spectral anomaly, formulated as follows:

$$\mathcal{H} = \mathcal{C} \times_3 B + \mathcal{E}_1,$$

The model represents $\mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ as a combination of three components: a spectral dictionary $B \in \mathbb{R}^{n_3 \times b}$, whose columns represent the primary spectral signatures of the background; a coefficient tensor $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times b}$, which encodes the spatial abundance of each background signature; and a spectral anomaly tensor $\mathcal{E}_1 \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

First, since spectral signatures are inherently non-negative, we impose a non-negativity constraint on the dictionary, $B \geq 0$. Second, to eliminate the inherent scale ambiguity between B and $\mathcal{C}$, we enforce a unit-norm constraint on each spectral tube of the coefficient tensor, i.e., $\|\mathcal{C}(i,j,:)\| = 1$.

$$\mathcal{H} = \mathcal{C} \times_3 B + \mathcal{E}_1, \quad B \geq 0,\ \|\mathcal{C}(i,j,:)\| = 1,\ \forall i \in [n_1],\ j \in [n_2].$$

We account for the presence of Gaussian noise in the data by using a Frobenius norm fidelity term. Furthermore, to enforce the prior knowledge that anomalies are spatially sparse yet spectrally dense, we penalize the anomaly tensor $\mathcal{E}_1$ using the joint sparsity inducing $\ell_{1,1,\phi}$-norm. Integrating all these components leads to the following objective function:

$$\begin{aligned} \min_{\mathcal{C},B,\mathcal{E}_1} \quad & \tfrac{\lambda_1}{2}\|B\|^2 + \lambda_2\|\mathcal{E}_1\|_{11\phi} + \tfrac{\lambda_3}{2}\|\mathcal{H} - \mathcal{C}\times_3 B - \mathcal{E}_1\|^2 \\ \text{s.t.} \quad & B \geq 0,\ \|\mathcal{C}(i,j,:)\| = 1,\ \forall i \in [n_1],\ j \in [n_2], \end{aligned} \tag{1}$$

where $\|\mathcal{E}_1\|_{11\phi} := \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\phi(\|\mathcal{E}_1(i,j,:)\|)$ with function $\phi(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$.

Layer 2: Spatial anomaly extraction

The coefficient tensor $\mathcal{C}$, produced by the spectral decomposition in the first layer, serves as the input for the second layer of our model. While the spectral dimension has been compressed, $\mathcal{C}$ inherits the strong spatial self-similarity of the background from the original HSI, resulting in a spatially low rank structure. This leads to the following optimization problem:

$$\begin{aligned} \min_{\mathcal{L},\mathcal{E}_2} \quad & \lambda_4 \operatorname{rank}_t(\mathcal{L}) + \lambda_5 \|\mathcal{E}_2\|_{11\phi} \\ \text{s.t.} \quad & \mathcal{C} = \mathcal{L} + \mathcal{E}_2. \end{aligned} \tag{2}$$

Here $\mathcal{L}$, $\mathcal{E}_2 \in \mathbb{R}^{n_1 \times n_2 \times b}$.

定理 1

For any tensor $\mathcal{L} \in \mathbb{R}^{n_1 \times n_2 \times b}$, *one has*

$$\operatorname{rank}_t(\mathcal{L}) = \min\left\{ \|\mathcal{Z}\|_{F,0} : \mathcal{L} = \mathcal{D} * \mathcal{Z}^T, \mathcal{D}^T * \mathcal{D} = \mathcal{I} \right\},$$

where $\|\mathcal{Z}\|_{F,0} = \sum_{j=1}^{r} \|\mathcal{Z}(:, j, :)\|^0$. *Here* $\mathcal{D} \in \mathbb{R}^{n_1 \times r \times b}$, $\mathcal{Z} \in \mathbb{R}^{n_2 \times r \times b}$, *and* $\operatorname{rank}_t(\mathcal{L}) \leq r \leq \min\{n_1, n_2\}$.

定理 2

There exists a $\bar{\nu} > 0$. *For any function* ψ *that satisfies* $\psi(0) = 0$, $x/\nu \leq \psi(x) < 1$ *for* $x \in (0, \nu)$ *and* $\psi(x) = 1$ *for* $x \geq \nu$, *where* $0 < \nu < \bar{\nu}$, *the following two problems share the same global minimizers and optimal values:*

$$\begin{cases} (P_0)\ \min \left\{ \|\mathcal{Z}\|_{F,0} : \mathcal{L} = \mathcal{D} * \mathcal{Z}^T, \mathcal{D}^T * \mathcal{D} = \mathcal{I} \right\}; \\ (P_\psi)\ \min \left\{ \|\mathcal{Z}\|_{F,1}^{\psi} : \mathcal{L} = \mathcal{D} * \mathcal{Z}^T, \mathcal{D}^T * \mathcal{D} = \mathcal{I} \right\}. \end{cases}$$

Here $\|\mathcal{Z}\|_{F,1}^{\psi} = \sum_{j=1}^{r} \psi(\|\mathcal{Z}(:, j, :)\|)$.

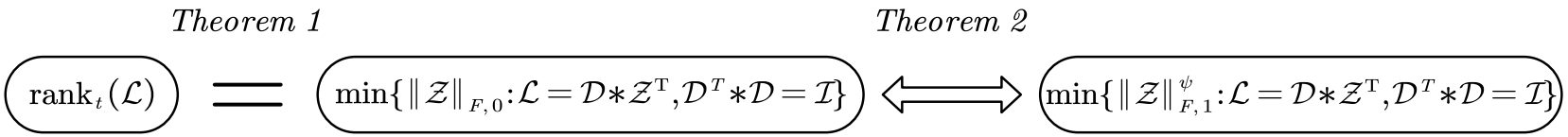


Figure 4: Relationships between tensor tubal rank and group sparsity regularization.

Motivated by Theorems 1-2, we transform Problem (2) into the following formulation:

$$\begin{aligned} \min_{\mathcal{D},\mathcal{Z},\mathcal{E}_2} \quad & \lambda_4 \|\mathcal{Z}\|_{F,1}^{\psi} + \lambda_5 \|\mathcal{E}_2\|_{11\phi} \\ \text{s.t.} \quad & \mathcal{C} = \mathcal{D} * \mathcal{Z}^T + \mathcal{E}_2,\ \mathcal{D}^T * \mathcal{D} = \mathcal{I}, \end{aligned} \tag{3}$$

where $\mathcal{E}_2 \in \mathbb{R}^{n_1 \times n_2 \times b}$, $\mathcal{D} \in \mathbb{R}^{n_1 \times r \times b}$, $\mathcal{Z} \in \mathbb{R}^{n_2 \times r \times b}$.

Remark 3.1

Optimizing $\|\mathcal{Z}\|_{F,1}^{\psi}$ is significantly more efficient than optimizing $\|\mathcal{L}\|_*$, as it eliminates the need for large-scale matrix SVD and the dimensionality of $\mathcal{Z}$ is considerably smaller than that of $\mathcal{L}$. Furthermore, in algorithms based on $\|\mathcal{Z}\|_{F,1}^{\psi}$, the size of r can adaptively decrease until it approximates $\mathrm{rank}_t(\mathcal{L})$, thereby obviating the need for additional techniques to adjust r.

LTD

Integrating the spectral analysis from Layer 1 and the spatial analysis from Layer 2, the proposed LTD model can be formulated as:

$$\begin{aligned} \min_{\mathcal{C},B,\mathcal{E}_1,\mathcal{D},\mathcal{Z},\mathcal{E}_2} \quad & \tfrac{\lambda_1}{2}\left\|B\right\|^2 + \lambda_2\left\|\mathcal{E}_1\right\|_{11\phi} + \tfrac{\lambda_3}{2}\left\|\mathcal{H} - \mathcal{C}\times_3 B - \mathcal{E}_1\right\|^2 \\ & +\lambda_4\left\|\mathcal{Z}\right\|_{F,1}^{\psi} + \lambda_5\left\|\mathcal{E}_2\right\|_{11\phi} \\ \text{s.t.} \quad & \mathcal{C} = \mathcal{D} * \mathcal{Z}^T + \mathcal{E}_2,\ B \in \mathbb{B},\ \mathcal{C} \in \mathbb{C},\ \mathcal{D} \in \mathbb{D}, \end{aligned} \tag{4}$$

where $\mathbb{B} = \{B \mid B \geq 0\}$, $\mathbb{C} = \{\mathcal{C} \mid \|\mathcal{C}(i,j,:)\| = 1\}$, and $\mathbb{D} = \left\{\mathcal{D} \mid \mathcal{D}^T * \mathcal{D} = I\right\}$.

Contents

By employing the half-quadratic splitting technique, we convert the constrained problem (4) into the subsequent unconstrained formulation:

$$\begin{aligned}\min \quad & F\left(\mathcal{C}, B, \mathcal{E}_1, \mathcal{D}, \mathcal{Z}, \mathcal{E}_2\right) = \frac{\lambda_1}{2}\left\|B\right\|^2 + \lambda_2\left\|\mathcal{E}_1\right\|_{11\phi} \\ & + \frac{\lambda_3}{2}\left\|\mathcal{H} - \mathcal{C} \times_3 B - \mathcal{E}_1\right\|^2 + \lambda_4\left\|\mathcal{Z}\right\|_{F,1}^{\psi} \\ & + \lambda_5\left\|\mathcal{E}_2\right\|_{11\phi} + \frac{\lambda_6}{2}\left\|\mathcal{C} - \mathcal{D} * \mathcal{Z}^T - \mathcal{E}_2\right\|^2 \\ & + \delta_{\mathbb{B}}\left(B\right) + \delta_{\mathbb{C}}\left(\mathcal{C}\right) + \delta_{\mathbb{D}}\left(\mathcal{D}\right).\end{aligned} \tag{5}$$

Within the framework of the PAM algorithm, we alternatively update each variable:

$$\begin{cases} \mathcal{C}^{t+1} = \arg\min_{\mathcal{C}} \ F\left(\mathcal{C}, B^t, \mathcal{E}_1^t, \mathcal{D}^t, \mathcal{Z}^t, \mathcal{E}_2^t\right) + \frac{\rho_1}{2}\left\|\mathcal{C} - \mathcal{C}^t\right\|^2, \\ B^{t+1} = \arg\min_{B} \ F\left(\mathcal{C}^{t+1}, B, \mathcal{E}_1^t, \mathcal{D}^t, \mathcal{Z}^t, \mathcal{E}_2^t\right) + \frac{\rho_2}{2}\left\|B - B^t\right\|^2, \\ \mathcal{E}_1^{t+1} = \arg\min_{\mathcal{E}_1} \ F\left(\mathcal{C}^{t+1}, B^{t+1}, \mathcal{E}_1, \mathcal{D}^t, \mathcal{Z}^t, \mathcal{E}_2^t\right) + \frac{\rho_3}{2}\left\|\mathcal{E}_1 - \mathcal{E}_1^t\right\|^2, \\ \mathcal{D}^{t+1} = \arg\min_{\mathcal{D}} \ F\left(\mathcal{C}^{t+1}, B^{t+1}, \mathcal{E}_1^{t+1}, \mathcal{D}, \mathcal{Z}^t, \mathcal{E}_2^t\right) + \frac{\rho_4}{2}\left\|\mathcal{D} - \mathcal{D}^t\right\|^2, \\ \mathcal{Z}^{t+1} = \arg\min_{\mathcal{Z}} \ F\left(\mathcal{C}^{t+1}, B^{t+1}, \mathcal{E}_1^{t+1}, \mathcal{D}^{t+1}, \mathcal{Z}, \mathcal{E}_2^t\right) + \frac{\rho_5}{2}\left\|\mathcal{Z} - \mathcal{Z}^t\right\|^2, \\ \mathcal{E}_2^{t+1} = \arg\min_{\mathcal{E}_2} \ F\left(\mathcal{C}^{t+1}, B^{t+1}, \mathcal{E}_1^{t+1}, \mathcal{D}^{t+1}, \mathcal{Z}^{t+1}, \mathcal{E}_2\right) + \frac{\rho_6}{2}\left\|\mathcal{E}_2 - \mathcal{E}_2^t\right\|^2, \end{cases}$$

where $\rho_i,\ i \in [6]$ are six positive constants, and t denotes the iteration number.

Spectral-spatial fusion

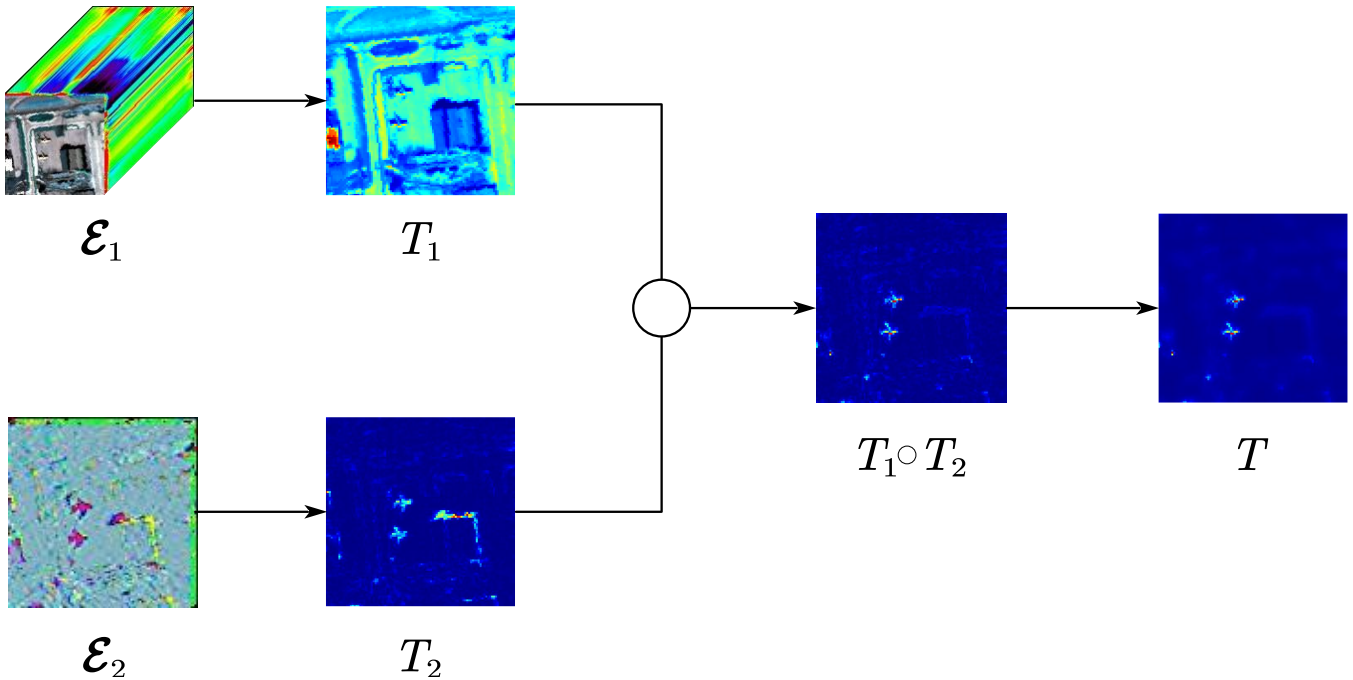


Figure 5: Flow chart of spectral-spatial fusion.

First, we compute the spectral anomaly detection map, given by $T_1 = \sqrt{\sum_{k=1}^{n_3} |\mathcal{E}_1^*(:,:,k)|^2}$, and the spatial anomaly detection map, given by $T_2 = \sqrt{\sum_{k=1}^{b} |\mathcal{E}_2^*(:,:,k)|^2}$. Then, we obtain the final detection result as follows:

$$T = \mathrm{IGF}\left(T_1 \circ T_2\right) \tag{6}$$

or

$$T = \mathrm{IGF}(\mathrm{IGF}(\mathrm{IGF}\left(T_1 \circ T_2\right), T_1), T_2), \tag{7}$$

where $\circ$ represents the Hadamard product operator, and IGF is a guided image filter.

Algorithm

Algorithm 1: The PAM-based solver for the proposed LTD model.

Input: HSI data $\mathcal{H}$, regularization parameters $\{\lambda_i\}_{i=1}^{6}$ and proximal $\{\rho_i\}_{i=1}^{6}$.

Output: Detection map T.

$t \leftarrow 0$;

while *not converged* **do**

  Update $\mathcal{C}^{t+1}$, B^{t+1}, $\mathcal{E}_1^{t+1}$, $\mathcal{D}^{t+1}$, $\mathcal{Z}^{t+1}$, $\mathcal{E}_2^{t+1}$ respectively;

  $t \leftarrow t + 1$;

end

Obtain the optimal anomalies $\mathcal{E}_1^*$, $\mathcal{E}_2^*$;

Compute the detection map T by (6) or (7).

Convergence analysis

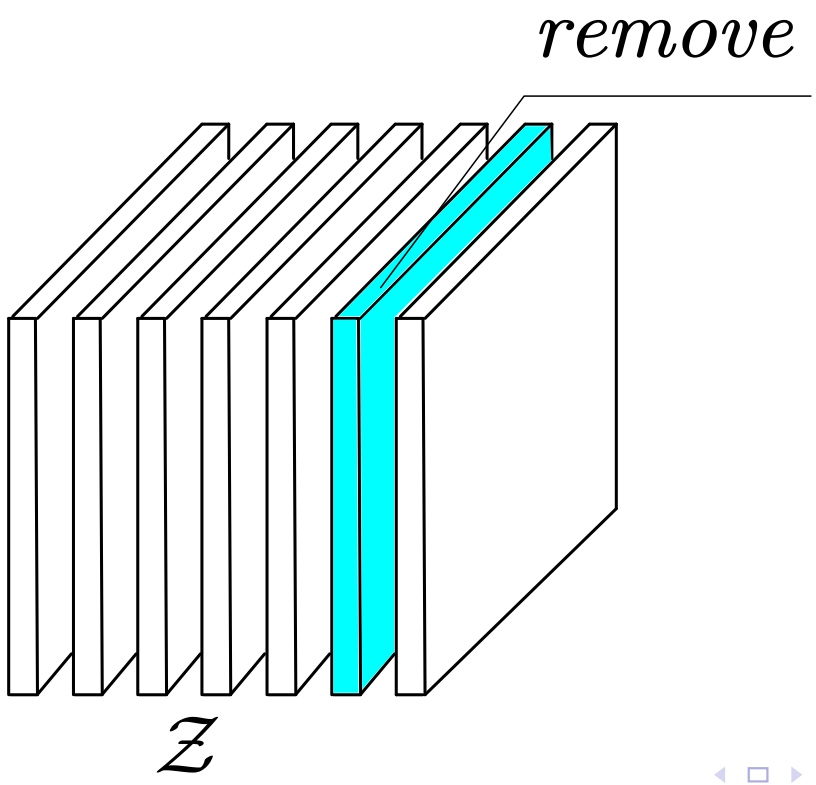
定理 3

Consider the sequence $\{\mathcal{W}^t\}$ obtained by Algorithm 1. Assuming that either $\mathcal{Z}^t$ or $\mathcal{E}_2^t$ is bounded, the sequence $\{\mathcal{W}^t\}$ converges to a critical point of $F(\mathcal{W})$.

Rank reduction strategy with validation mechanism

定理 4

There exists $t^{\#} \in \mathbb{N}$ *such that* $\Gamma(\mathcal{Z}^t) = \Gamma(\mathcal{Z}^{t+1})$ *for each* $t \geq t^{\#}$, *where* $\Gamma(\mathcal{Z}) = \{j \mid \|\mathcal{Z}(:, j, :)\| \neq 0,\ j = 1, \ldots, r\}$.



Algorithm 2: Rank reduction with validation mechanism strategy.

Input: HSI data $\mathcal{H}$, parameters $\{\lambda_i\}_{i=1}^6$ and $\{\rho_i\}_{i=1}^6$. Set $\mathbb{S} = \emptyset$.
Output: Detection map T.

$t \leftarrow 0$;
while *not converged* **do**
  Update $\mathcal{C}^{t+1}$, B^{t+1}, $\mathcal{E}_1^{t+1}$, $\mathcal{D}^{t+1}$, $\mathcal{Z}^{t+1}$, and $\mathcal{E}_2^{t+1}$;
  if $\mathbb{S} \neq \emptyset$ **then**
    Compute

$$\mathcal{Z}_{sub}^{t+1} = \mathrm{Prox}_{\hat{\lambda}_4}\, \psi\left(\lambda_6 \left(\mathcal{C}^{t+1} - \mathcal{E}_2^{t+1}\right)^T * \mathcal{D}_{sub}^t(:, \mathbb{S}, :)/(\lambda_6 + \rho_5)\right);$$

    Identify the top 5 lateral slices of $\mathcal{Z}_{sub}^{t+1}$ with the highest Frobenius norms, or all if fewer than 5, and denote this set as $\mathbb{S}_1$;
    Let $\mathcal{Z}^{t+1} = [\mathcal{Z}^{t+1}, \mathcal{Z}_{sub}^{t+1}(:, \mathbb{S}_1, :)]$ and $\mathcal{D}^{t+1} = [\mathcal{D}^{t+1}, \mathcal{D}_{sub}^t(:, \mathbb{S}_1, :)]$;
  end
  Let $\mathbb{S} = \{j \mid \|\mathcal{Z}^{t+1}(:, j, :)\| = 0\}$, and define $\mathcal{D}_{sub}^{t+1} = \mathcal{D}^{t+1}(:, \mathbb{S}, :)$. Then, remove the lateral slices at positions in $\mathbb{S}$ from $\mathcal{D}^{t+1}$ and $\mathcal{Z}^{t+1}$;
  $t \leftarrow t + 1$;
end
Obtain the optimal anomalies $\mathcal{E}_1^*$, $\mathcal{E}_2^*$;
Compute the detection map T by (6) or (7).

Contents

Contents

Effects of rank reduction strategy with validation mechanism

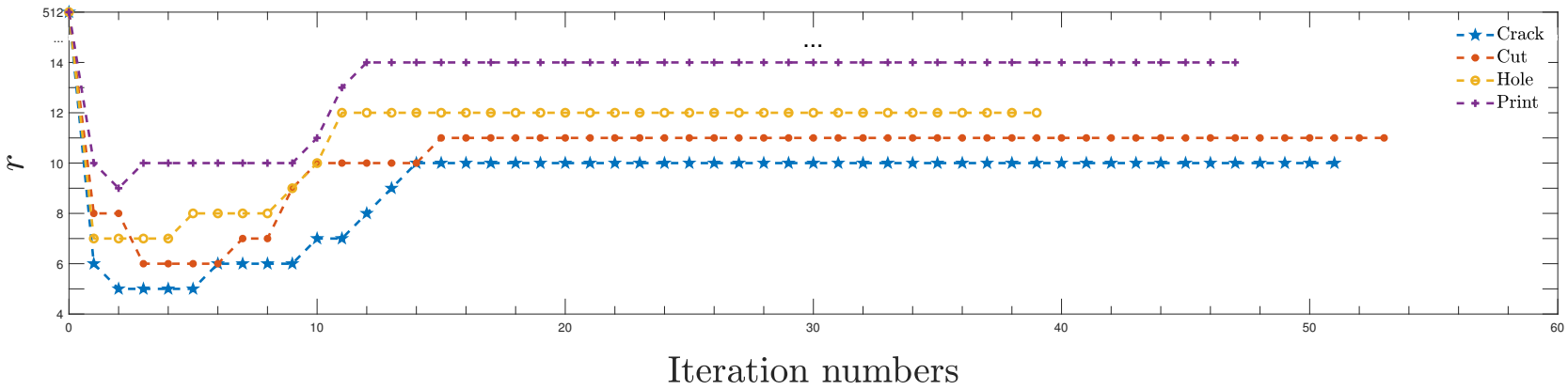


Figure 6: Variation of r values across iterations for Algorithm 2.

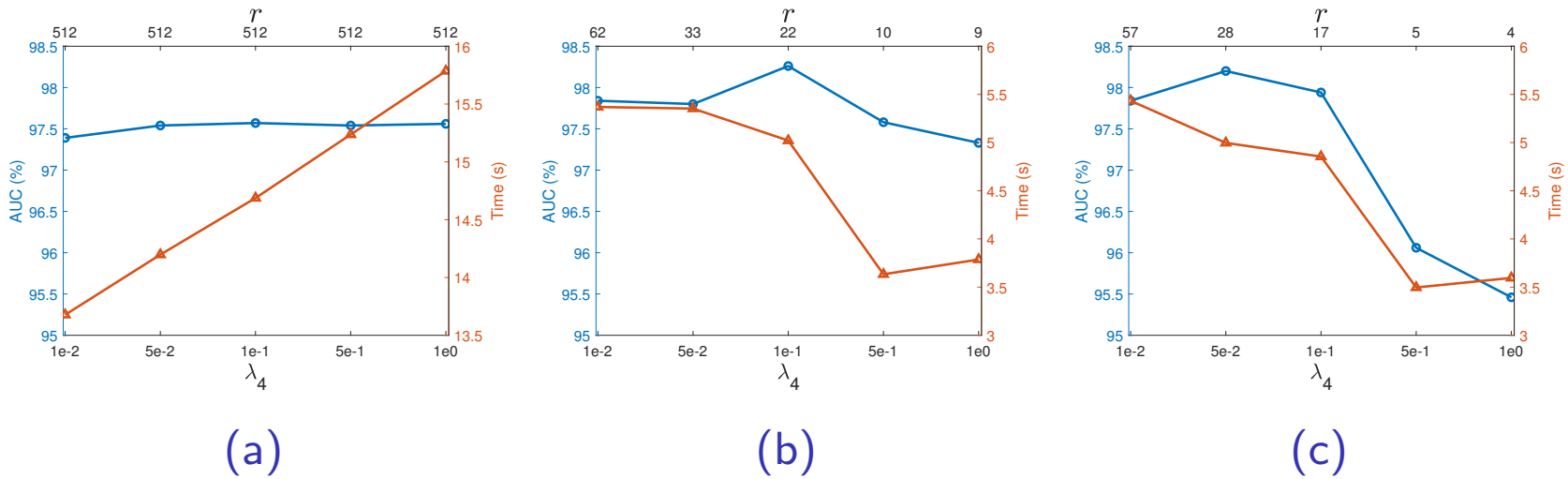
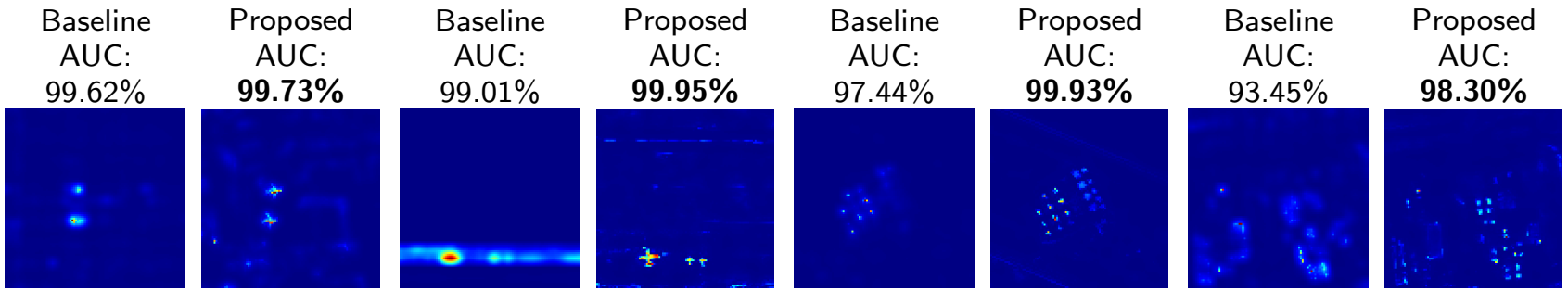


Figure 7: Performance analysis of (a) Algorithm 1, (b) Algorithm 2, and (c) Algorithm 2 without validation mechanism on AUC values (%), running time (s), and r metrics across different various values of λ_4 on the Crack dataset.

Effects of the unified framework

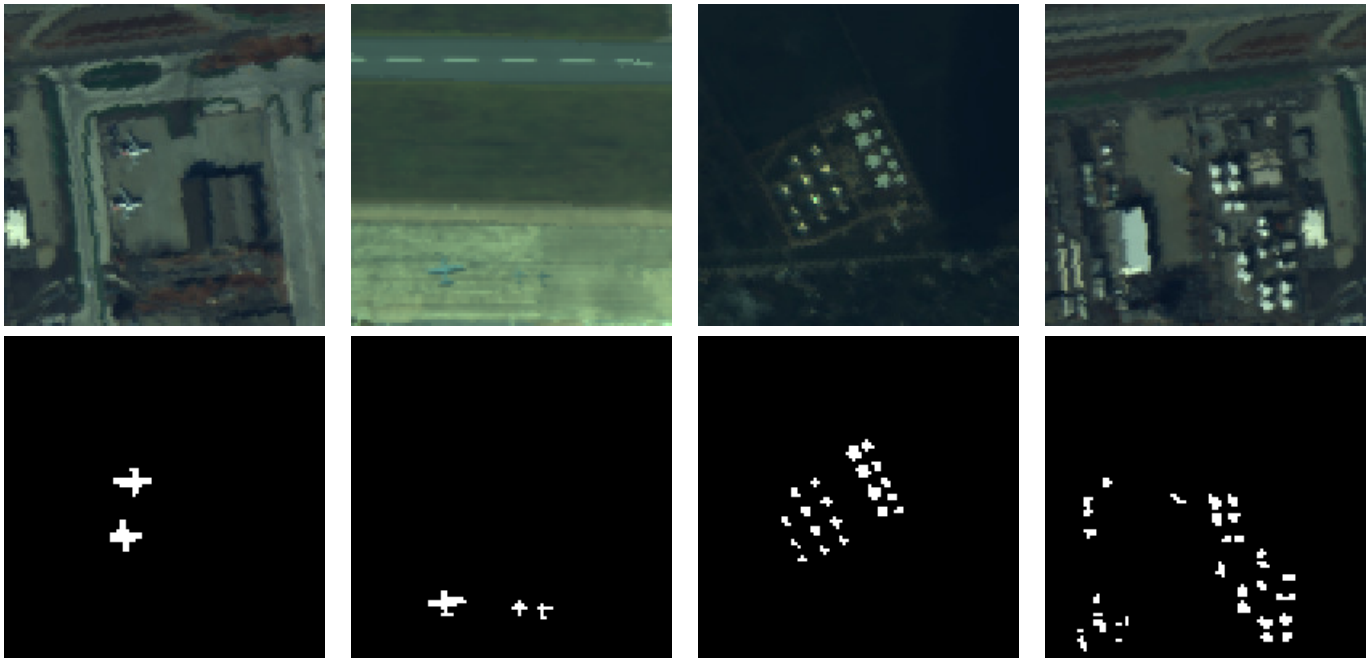


(a) Airport1 (b) Airport2 (c) Urban1 (d) Urban2

Figure 8: Target detection results for ABU dataset. For each scene, the left map is the result of the Baseline (Layer1+Layer2) and the right map is from our Proposed (LTD) method. The corresponding AUC score is displayed above each map.

Contents

ABU data set



(a) Airport1 (b) Airport2 (c) Urban1 (d) Urban2

Figure 9: Pseudo-color images and ground-truth maps of ABU data sets.

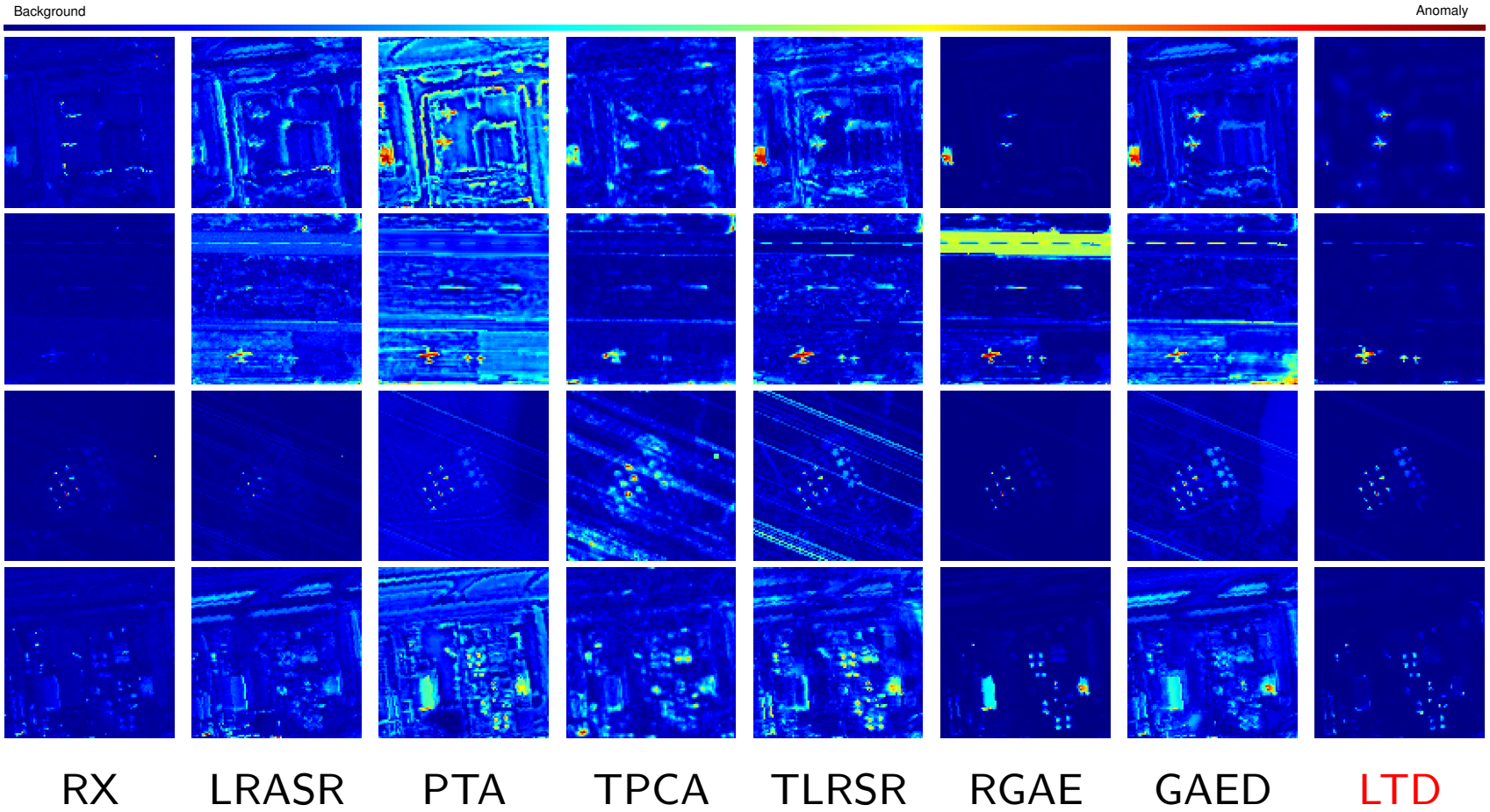
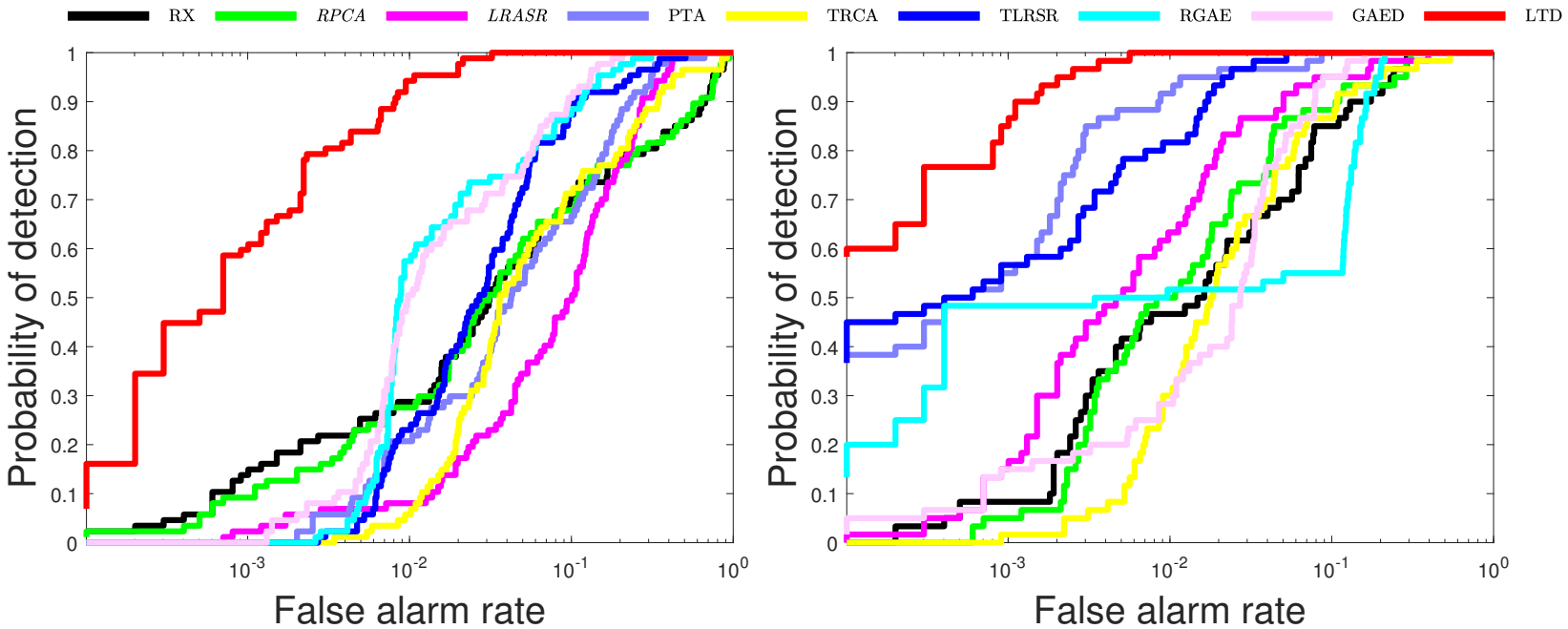


Figure 10: Target detection results by different methods for ABU data sets.

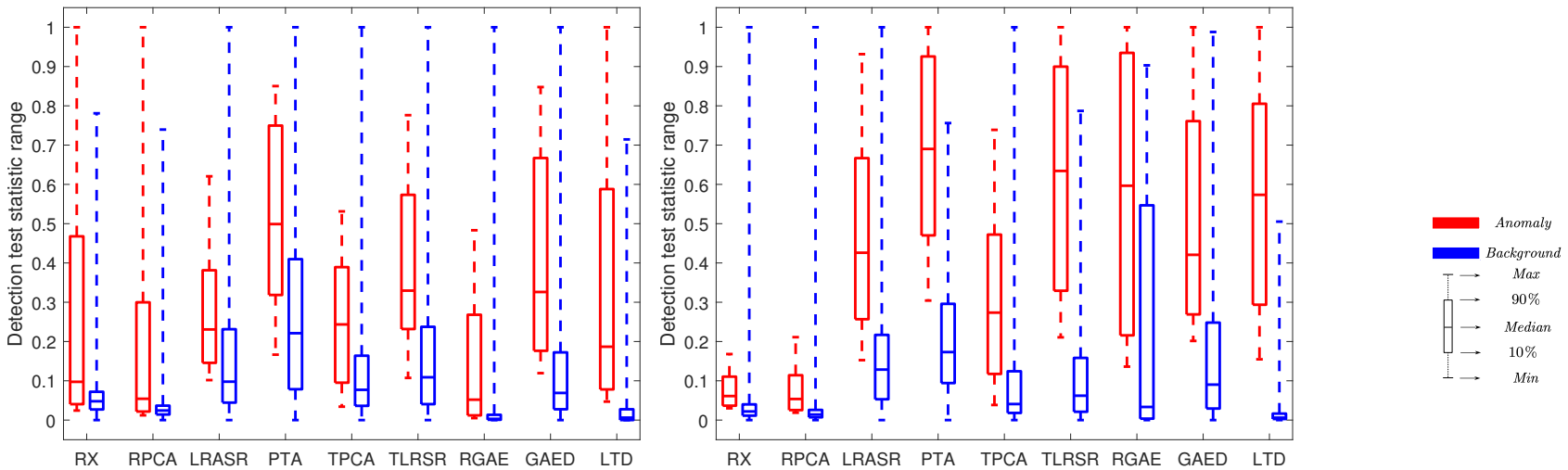
Table 1: Comparison of AUC values (%) and running time (s) of different methods for ABU dataset.

Algorithm	Index	RX	RPCA	LRASR	PTA	TPCA	TLRSR	RGAE	GAED	LTD
Airport1	AUC	84.04	84.28	87.70	90.96	88.11	94.58	96.40	96.77	**99.73**
	Time	0.05	4.62	20.06	15.77	18.63	2.50	51.88	41.10	1.66
Airport2	AUC	95.26	96.27	97.95	99.55	95.26	99.49	93.25	96.81	**99.95**
	Time	0.06	1.93	19.45	14.68	17.93	2.28	50.00	39.15	1.89
Urban1	AUC	99.46	99.57	87.03	97.70	95.45	95.42	99.73	99.59	**99.93**
	Time	0.05	3.98	22.87	16.69	17.98	2.40	53.78	39.84	2.13
Urban2	AUC	96.92	96.58	89.13	82.58	93.00	97.11	94.90	90.18	**98.30**
	Time	0.06	3.85	20.30	15.99	19.32	2.49	42.26	37.99	1.89



(a) Airport1 (b) Airport2

Figure 11: ROC curves obtained by different methods for ABU dataset.



(a) Airport1 (b) Airport2

Figure 12: Separability maps of different methods for ABU dataset.

MVTec data set

(a) Crack (b) Cut (c) Hole (d) Print

Figure 13: Color images and ground-truth maps of MVTec data sets.

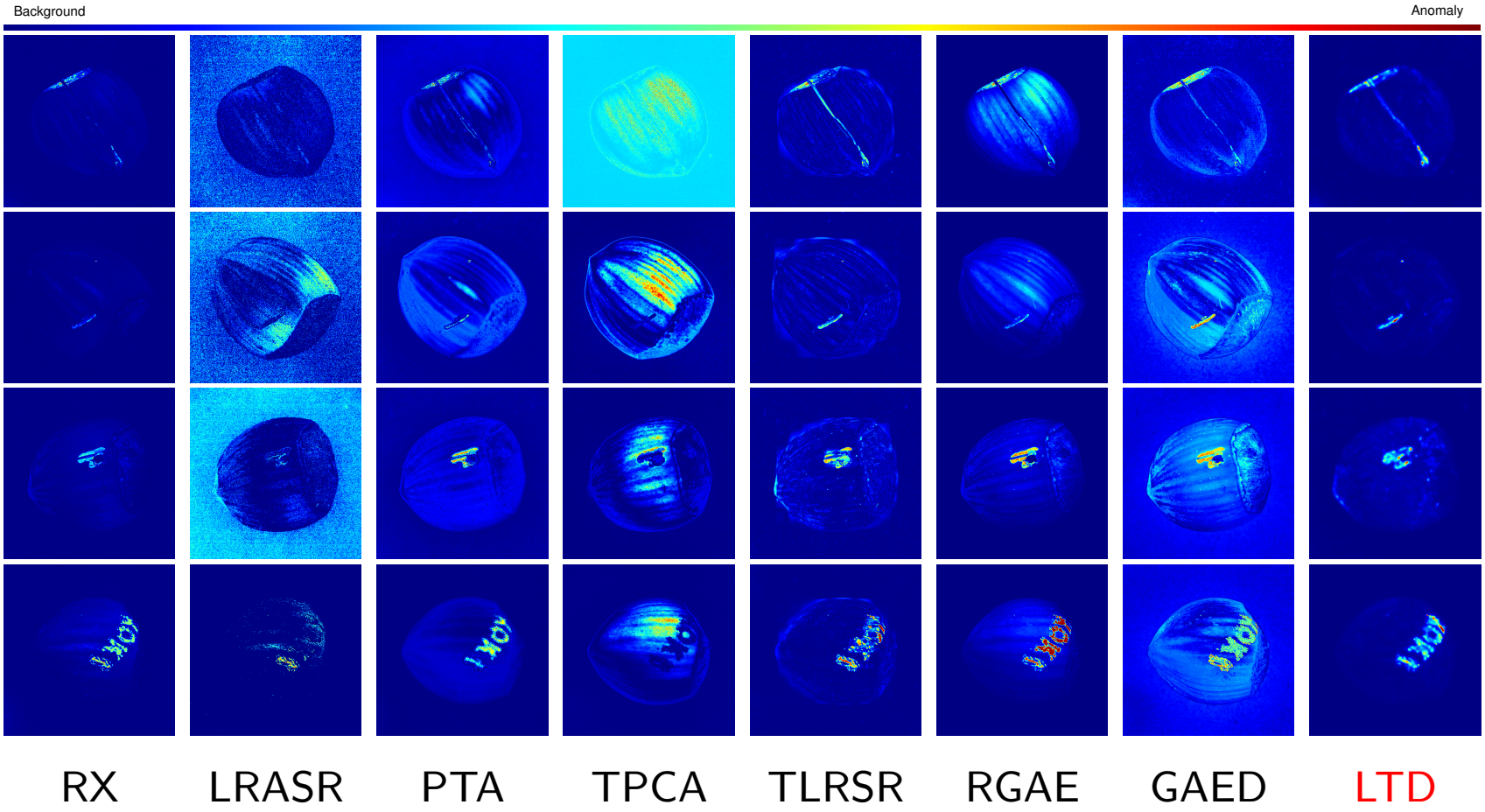
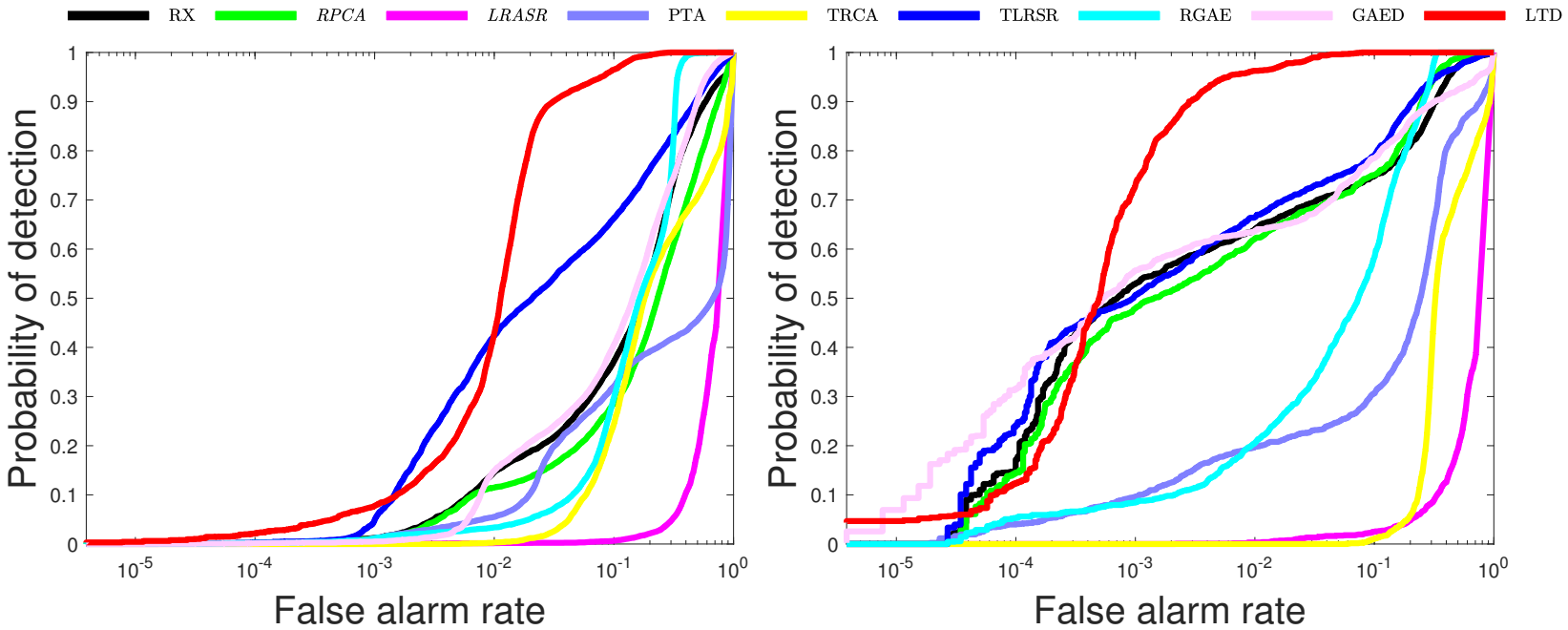


Figure 14: Target detection results by different methods for MVTec data sets.

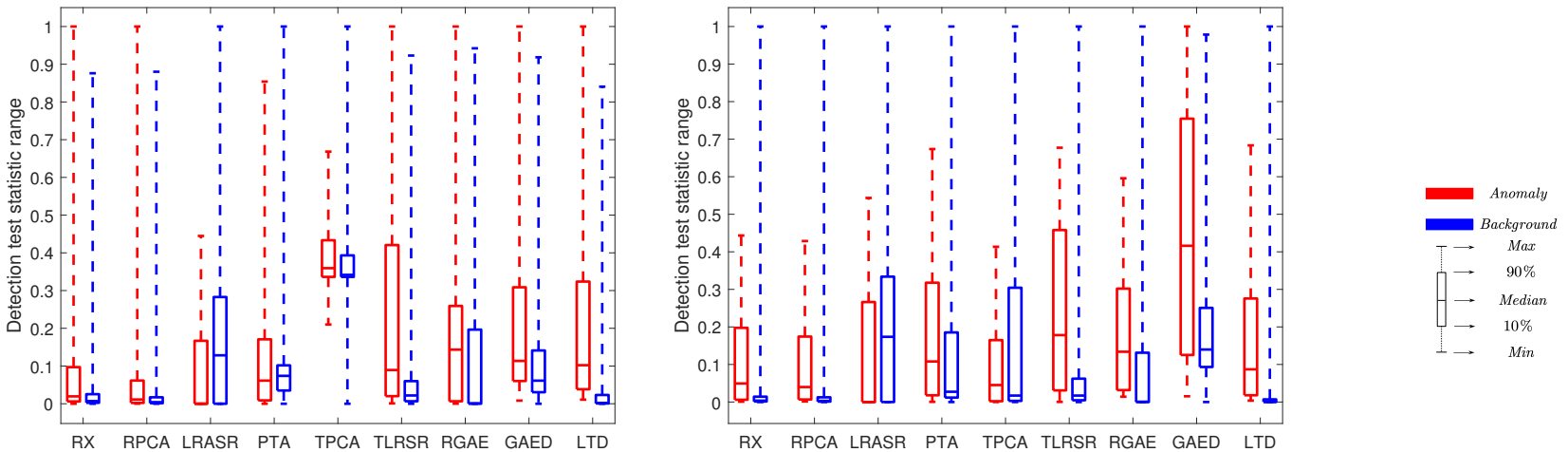
Table 2: Comparison of AUC values (%) and running time (s) of different methods for MVTec dataset.

Algorithm	Index	RX	RPCA	LRASR	PTA	TPCA	TLRSR	RGAE	GAED	LTD
Crack	AUC	77.23	71.12	29.98	45.66	65.75	83.34	82.03	78.37	**98.26**
	Time	0.05	0.40	977.33	16.04	14.33	29.41	818.99	669.92	4.99
Cut	AUC	89.46	91.50	33.48	72.30	56.40	91.11	88.54	86.02	**99.76**
	Time	0.05	0.40	991.03	16.81	13.96	32.88	837.77	662.67	4.75
Hole	AUC	91.26	88.50	22.93	78.77	82.01	95.75	93.88	90.34	**99.25**
	Time	0.05	0.40	979.06	16.26	14.62	46.29	814.96	677.04	2.62
Print	AUC	95.70	95.60	58.68	89.03	68.48	94.21	97.05	87.42	**99.87**
	Time	0.04	0.40	982.01	15.31	14.00	32.08	978.39	684.22	3.44



(a) Crack (b) Cut

Figure 15: ROC curves obtained by different methods for MVTec data sets.



(a) Crack (b) Cut

Figure 16: Separability maps of different methods for MVTec data sets.

Contents

Our contributions

The fundamental novelty of our LTD framework lies in two key areas: (1) a unified framework for enhanced performance and (2) specific architectural innovations for superior efficiency.

- First, to the best of our knowledge, LTD is the first framework based on low rank representation (LRR) to unify the detection of both spectral $\mathcal{E}_1$ and spatial $\mathcal{E}_2$ anomalies within a single optimization procedure. Unlike a simple sequential process, which completes spectral detection in Layer 1 before independently conducting spatial analysis in Layer 2, our unified optimization enables mutual refinement between the two layers. This synergy leads to significantly more accurate and robust detection performance, as empirically demonstrated in our ablation study.

- Second, the framework is explicitly designed for computational efficiency through two primary contributions. (a) The proposed layered decomposition structurally breaks down the large-scale input tensor into multiple, smaller tensors, significantly reducing the complexity of the optimization problem.

(b) Crucially, we establish the relationship between tensor tubal rank and group sparsity regularization in Theorems 1 and 2. Leveraging this insight, we employ group sparsity regularization in Layer 2 to characterize the spatial low-rankness of the tensor $\mathcal{C}$, rather than relying on conventional nuclear function regularization. Although skinny T-SVD computation is required for $\mathcal{D} \in \mathbb{R}^{n_1 \times r \times b}$, it is limited to small-scale tensors, as our algorithm adaptively reduces the size of $\mathcal{D}$. Consequently, the final r is much smaller than $\min\{n_1, n_2\}$. Thus, the proposed group sparsity regularization is markedly more efficient than nuclear function regularization.

Thanks for your attention!