

Supplementary Materials for “Hyperspectral Anomaly Detection Fused Unified Nonconvex Tensor Ring Factors Regularization ”

Wenjin Qin, Hailin Wang, *Student Member, IEEE*, Hao Shu, Feng Zhang, Jianjun Wang, *Member, IEEE*, Xiangyong Cao, *Member, IEEE*, Xi-Le Zhao, and Gemine Vivone, *Senior Member, IEEE*

In this supplementary material, the convergence theory of the proposed algorithm is presented in Section I. We provide parameters setting of the proposed and compared HAD algorithms in Section II. More experimental results regarding empirical comparison, parameter analysis, and ablation study are presented in Section III-IV. More experimental results on larger-scale HSI datasets are displayed in Section V.

I. CONVERGENCE THEORY AND RELATED PROOFS

Below, we provide the convergence analysis of the proposed nonconvex HAD algorithm. The main results are given in Theorem I.1 and Theorem I.2 below. Before that, we present some lemmas. Due to limitations on the length of the Supplementary Material page, the detailed proofs of relevant theories and lemmas are provided in author’s Github: <https://github.com/Qinwenjinswu/QWJSWU-Convergence-Theory-Proof>.

Lemma I.1. *The sequences $\{\mathbf{Y}^{\{\nu\}}\}$ and $\{\mathbf{Q}^{(n,k)\{\nu\}}\}$ ($n, k = 1, 2, 3$) generated by the proposed nonconvex HAD Algorithm are bounded.*

Lemma I.2. *Suppose that the sequences $\{\mathbf{Y}^{\{\nu\}}\}$ and $\{\mathbf{Q}^{(n,k)\{\nu\}}\}$ ($n, k = 1, 2, 3$) generated by the proposed HAD Algorithm are bounded, then the sequences $\{(\mathcal{L}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{S}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{G}^{(n)})^{\{\nu\}}\}$, and $\mathcal{E}^{\{\nu\}}$ are bounded.*

Theorem I.1. *Suppose that the sequences $\{\mathbf{Y}^{\{\nu\}}\}$ and $\{\mathbf{Q}^{(n,k)\{\nu\}}\}$ ($n, k = 1, 2, 3$) generated by the proposed HAD Algorithm are bounded. Then, the sequences $\{(\mathcal{L}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{S}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{G}^{(n)})^{\{\nu\}}\}$, and $\mathcal{E}^{\{\nu\}}$ satisfy:*

- 1) $\lim_{\nu \rightarrow \infty} \|\mathcal{M} - \mathfrak{R}(\mathcal{G}\mathcal{I}) - \mathcal{E}^{\{\nu+1\}}\|_F = 0;$
- 2) $\lim_{\nu \rightarrow \infty} \|\nabla_k (\mathcal{G}^{(n)})^{\{\nu+1\}} - (\mathcal{L}^{(n,k)})^{\{\nu+1\}} - (\mathcal{S}^{(n,k)})^{\{\nu+1\}}\|_F = 0, \quad n, k = 1, 2, 3;$
- 3) $\lim_{\nu \rightarrow \infty} \|(\mathcal{L}^{(n,k)})^{\{\nu+1\}} - (\mathcal{L}^{(n,k)})^{\{\nu\}}\|_F = \lim_{\nu \rightarrow \infty} \|\mathcal{E}^{\{\nu+1\}} - \mathcal{E}^{\{\nu\}}\|_F = 0;$
- 4) $\lim_{\nu \rightarrow \infty} \|(\mathcal{S}^{(n,k)})^{\{\nu+1\}} - (\mathcal{S}^{(n,k)})^{\{\nu\}}\|_F = \lim_{\nu \rightarrow \infty} \|(\mathcal{G}^{(n)})^{\{\nu+1\}} - (\mathcal{G}^{(n)})^{\{\nu\}}\|_F = 0; \quad n, k = 1, 2, 3.$

Theorem I.2. *Let $\{\mathbf{Y}^{\{\nu\}}\}$, $\{\mathbf{Q}^{(n,k)\{\nu\}}\}$ ($n, k = 1, 2, 3$), $\{(\mathcal{L}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{S}^{(n,k)})^{\{\nu\}}\}$, $\{(\mathcal{G}^{(n)})^{\{\nu\}}\}$, and $\mathcal{E}^{\{\nu\}}$ be the sequences generated by the proposed nonconvex HAD Algorithm. Suppose that the sequences $\{\mathbf{Y}^{\{\nu\}}\}$ and $\{\mathbf{Q}^{(n,k)\{\nu\}}\}$ ($n, k = 1, 2, 3$) are bounded. Then, any accumulation point of the sequence $\left\{ \{\mathbf{Y}^{\{\nu\}}\}, \{\mathbf{Q}^{(n,k)\{\nu\}}\} \right. (n, k = 1, 2, 3), \{(\mathcal{L}^{(n,k)})^{\{\nu\}}\}, \{(\mathcal{S}^{(n,k)})^{\{\nu\}}\}, \{(\mathcal{G}^{(n)})^{\{\nu\}}\}, \mathcal{E}^{\{\nu\}} \right\}$ is a Karush-Kuhn-Tucker (KKT) point of the original optimization problem.*

This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFA1008502; in part by the National Natural Science Foundation of China’s Regional Innovation Development Joint Fund under Grant U24A2001; in part by the Natural Science Foundation of Chongqing, China, under Grant CSTB2023NSCQ-LZX0044; in part by the National Natural Science Foundation of China under Grant 12301594, Grant 12201505, Grant 12101512; in part by the Chongqing Talent Project, China, under Grant csc2021ycjh-bgzm0015; in part by the Fundamental Research Funds for the Central Universities under Grant SWU-KR25013; and in part by the Initiative Projects for Ph.D. in China West Normal University under Grant 22kE030. (Corresponding author: Jianjun Wang.)

Wenjin Qin, Feng Zhang, and Jianjun Wang are with the School of Mathematics and Statistics, Southwest University, Chongqing 400715, China (e-mail: qinwenjin2021@163.com, zfm@swu.edu.cn, wjj@swu.edu.cn).

Hailin Wang and Hao Shu are with the School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an 710049, China (e-mail: wang-hailin97@163.com, haoshu812@gmail.com).

Xiangyong Cao is with School of Computer Science and Technology and Ministry of Education Key Lab For Intelligent Networks and Network Security, Xi’an Jiaotong University, Xi’an 710049, China (e-mail: caoxiangyong@mail.xjtu.edu.cn).

Xi-Le Zhao is with the School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: xlzhao122003@163.com).

Gemine Vivone is with the Institute of Methodologies for Environmental Analysis, CNR-IMAA, 85050 Tito Scalo, Italy, and also with the National Biodiversity Future Center (NBFC), 90133 Palermo, Italy (e-mail: gemine.vivone@imaa.cnr.it).

II. EXPERIMENTAL SETTINGS OF THE PROPOSED AND COMPARED METHODS

To maintain the fairness of our experiments, optimal parameters setting of all competing methods are given according to the relevant references [1]–[12]. Specifically, **RX** detector is parameter-free. The parameters of window size (w_{in}, w_{out}) in **CRD** algorithm are selected from $\{(3, 5), (5, 7), (7, 9), (9, 11), (11, 13), (13, 15), (15, 17), (17, 19), (19, 21)\}$. Besides, the regularization parameter λ is chosen from $\{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$. With regard to three deep learning-based methods, i.e., **GAED**, **PDBSNet**, and **GT-HAD** algorithms, their related parameters are mainly based on the references [3], [13]–[15]. For **LSMAD** algorithm, the empirical rank value is taken as $5\sim20$. Two parameters β and λ in **LRASR** algorithm are chosen from $[0.01, 0.05, 0.1, 0.5, 1, 5, 10]$. There are three regularization parameters in **GTVLRR** method, that is, λ , β , and γ . We determine the value of λ and β from $\{0.005, 0.05, 0.1, 0.3, 0.5, 0.7, 1, 5, 10\}$ and γ from $\{0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5\}$. For the background dictionary in **LRASR** and **GTVLRR** algorithms, the number of clusters is set to 15, and 20 pixels are selected in each cluster as atoms.

In the **PTA** method, the truncated rank value r is selected from $0\sim20$ and the regularization parameter μ is selected from $\{10^{-4}, 10^{-3}, 10^{-2}, 0.1, 1, 10, 10^2, 10^3, 10^4\}$. Other parameters of the **PTA** algorithm remain unchanged as in [7]. For **PCA-TLRSR** detector, the parameters λ and λ' are chosen from $\{0.0006, 0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$. The trade-off parameter λ in **TCTV** algorithm is set to be $\lambda=a/\sqrt{\min(n_1, n_2) \cdot n_3}$, a is selected from $\{1, 1.2, 1.5, 1.8, 2, 2.2\}$. The optimal combinations of two parameters λ and β in **TRDFTVAD**, are both searched from $\{0.003, 0.01, 0.03, 0.05, 0.1, 0.5, 1, 2, 5, 10, 15\}$. The hyperparameters θ and α involved in **GNLTR** method are selected from $\{0.0005, 0.001, 0.01, 0.03, 0.05, 0.07, 0.1\}$, and the non-convex surrogate is chosen as ℓ_q -norm. For the **GNBRL** detector, we set $\alpha_1, \alpha_2, \alpha_3 = \frac{1}{3}, \beta^0 \in \{0.001, 0.005, 0.01, 0.05, 0.1\}$, $\kappa \in \{1.3, 1.5, 1.7, 1.9\}$, $\lambda_1, \lambda_2 \in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1\}$. Besides, the nonconvex function ψ is set as Capped-Lp. Note that PCA technique is adopted to preprocess the original HSI data in **PCA-TLRSR**, **GNLTR** and **GNBRL** methods. For the proposed HAD method, the nonconvex combinations $\Phi(\cdot)+\psi(\cdot)$ are set to be the same ($\Phi=\psi=\text{Capped-Log}$), with the exception of Hyperion ($\Phi=\psi=\text{MCP}$), San-Diego ($\Phi=\psi=\text{Capped-Lp}$) and Urban-3 ($\Phi=\psi=\text{MCP}$). For Pavia, Airport-4, and Urban-5 datasets, the TR rank (r_1, r_2, r_3) is set to be $(6, 6, 6)$. For other remaining datasets, we assign $(r_1, r_2, r_3) = (6, 16, 6)$. The trade-off parameters α and β are both searched from $\{10^{-6}, 5 \cdot 10^{-6}, 10^{-5}, 5 \cdot 10^{-5}, 10^{-4}, 5 \cdot 10^{-4}, 10^{-3}, 5 \cdot 10^{-3}, 10^{-2}, 5 \cdot 10^{-2}, 0.1, 0.5, 1, 5, 10\}$, set $\mu^{\{0\}} = 10^{-3}$, $\mu^{\max} = 10^{10}$, $\vartheta = 1.1$, $\varpi = 10^{-5}$, $\nu_{\max} = 500$, $\mathcal{L} = \text{FFT}$.

III. MORE EMPIRICAL RESULTS ON COMPARED EXPERIMENTS AND PARAMETER ANALYSIS

More empirical results regarding compared experiments and parameter analysis are presented in this section.

- Figure 1 presents the anomaly detection maps of various HAD methods on other HSI datasets (HYDICE, Urban-5, Urban-3, Urban-4, Pavia, Beach-4);
- 3-D and 2-D ROC curves' performance comparison of different HAD approaches on other HSI datasets (HYDICE, Pavia, and San-Diego) are shown in Figure 2;
- Figure 3–6 investigates the impact of different trade-off parameters α and β upon detection performance under various nonconvex combinations $\Phi(\cdot)+\psi(\cdot)$, i.e., MCP+MCP, Capped-MCP+Capped-MCP, Lp+Lp, Capped-Lp+Capped-Lp.

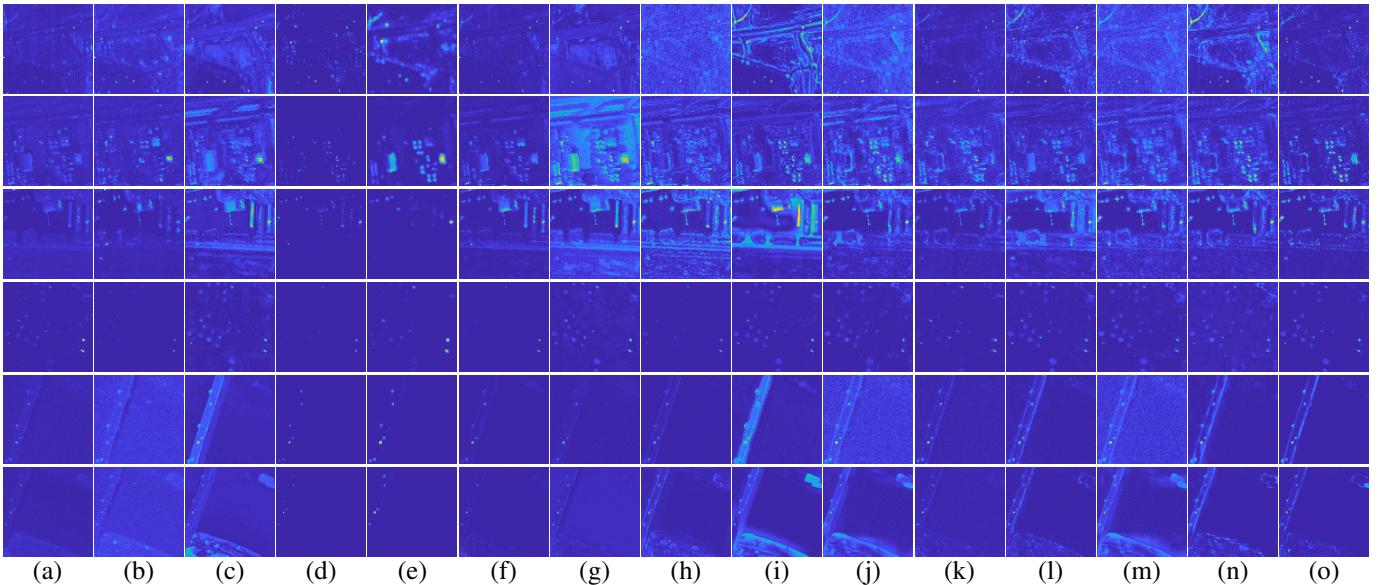


Fig. 1: Anomaly detection map of various HAD methods on six HSI datasets: HYDICE, Urban-5, Urban-3, Urban-4, Pavia, Beach-4. (from top to bottom). (a) RX. (b) CRD. (c) GAED. (d) PDBSNet. (e) GT-HAD. (f) LSMAD. (g) LRASR. (h) GTVLRR. (i) PTA. (j) PCA-TLRSR. (k) T-CTV. (l) TRDFTVAD. (m) GNLTR. (n) GNBRL. (o) Proposed.

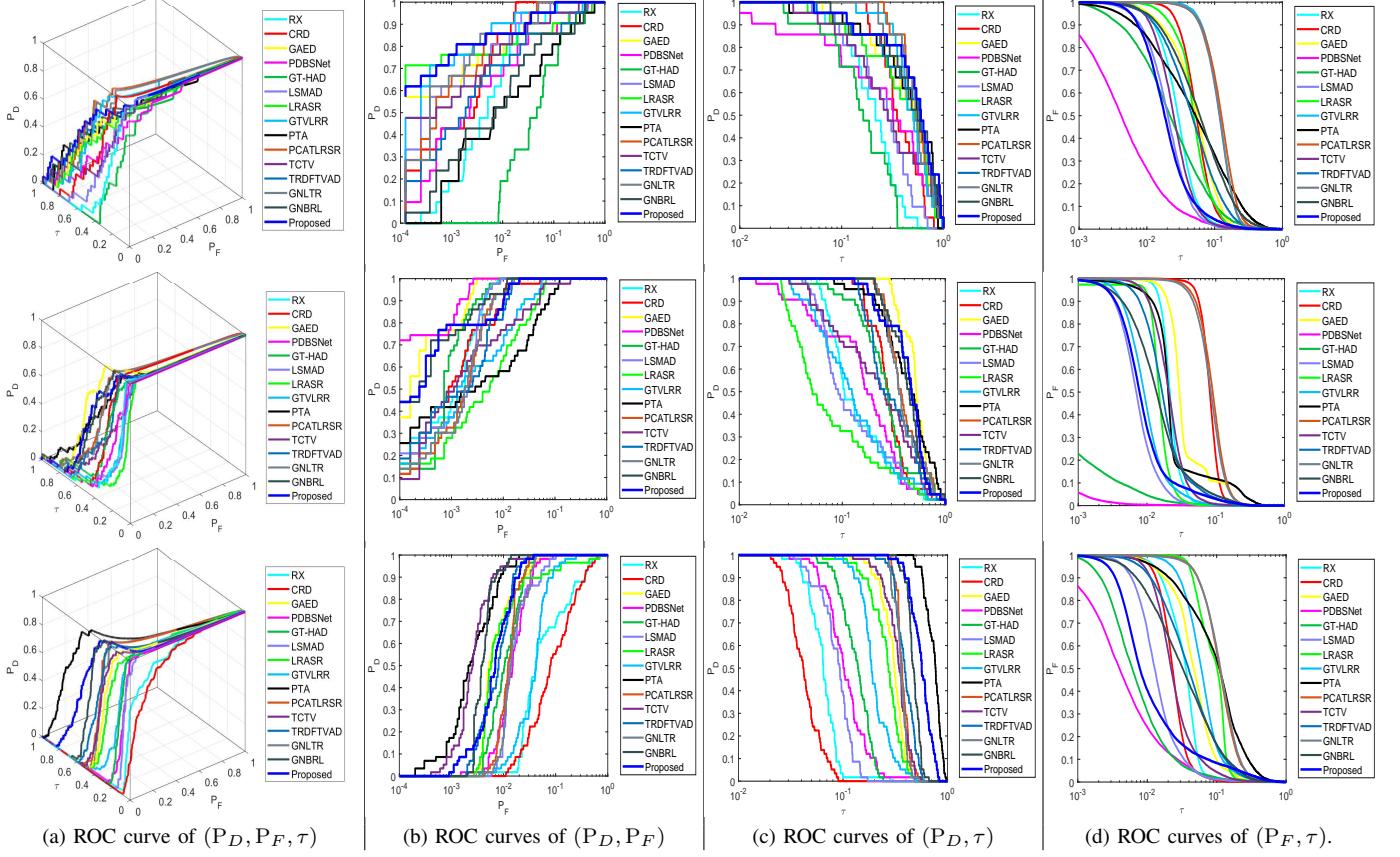


Fig. 2: 3-D and 2-D ROC curves' performance comparison of different HAD approaches on three HSI datasets: HYDICE, Pavia, and San-Diego (from top to bottom).

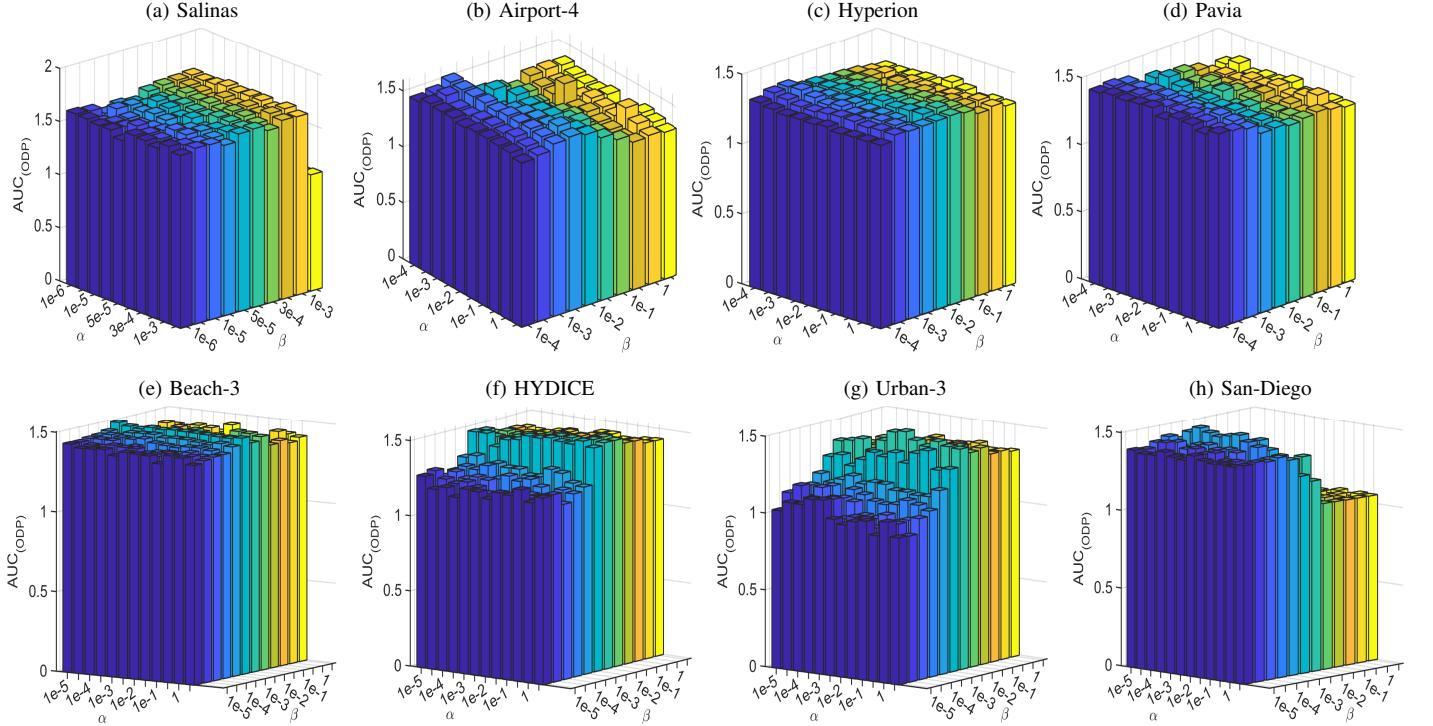


Fig. 3: Sensitivity analysis of the trade-off parameters α and β . For each HSI data, the TR rank is fixed to $(r_1, r_2, r_3) = (6, 6, 6)$. All nonconvex combinations $\Phi+\psi$ are set to be MCP+MCP.

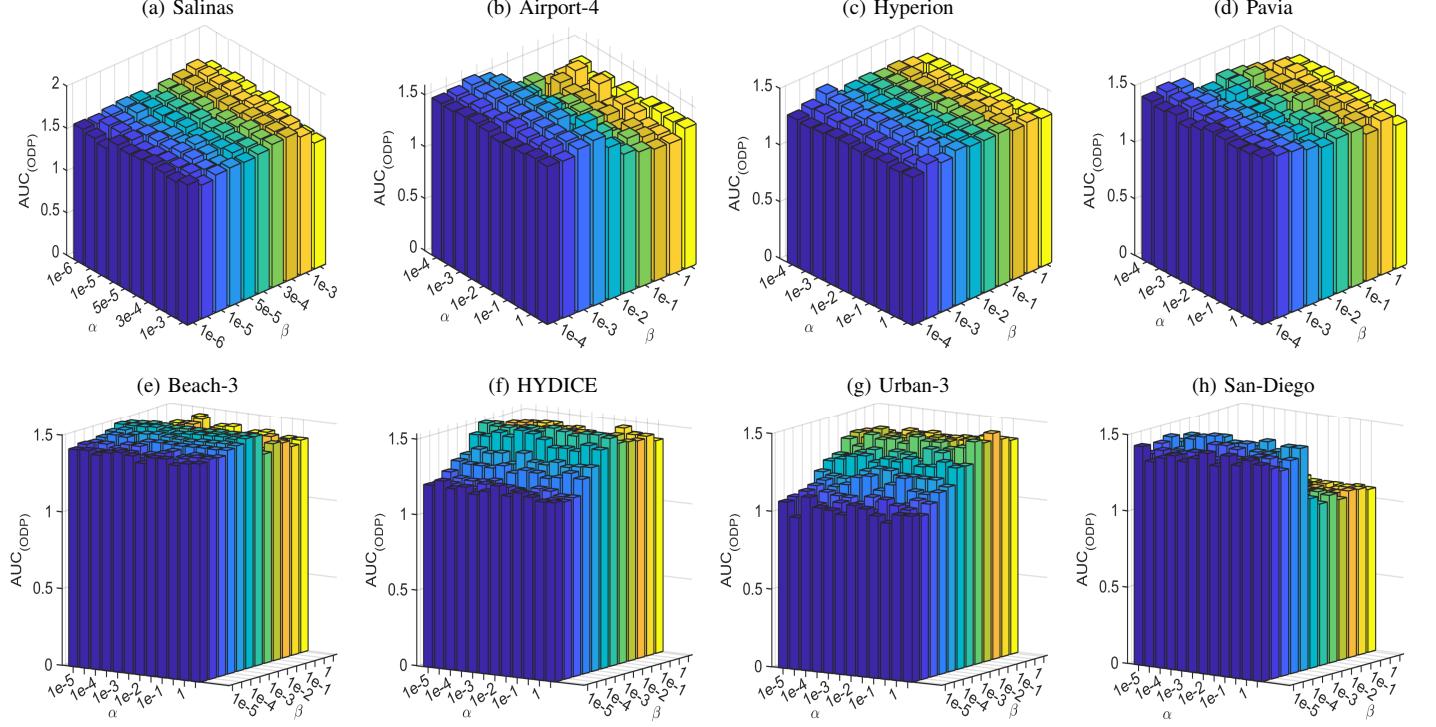


Fig. 4: Sensitivity analysis of the trade-off parameters α and β . For each HSI data, the TR rank is fixed to $(r_1, r_2, r_3) = (6, 6, 6)$, and the nonconvex combination $\Phi+\psi$ is set to be the same, i.e., $\Phi(\cdot) = \psi(\cdot)$ =Capped-MCP.

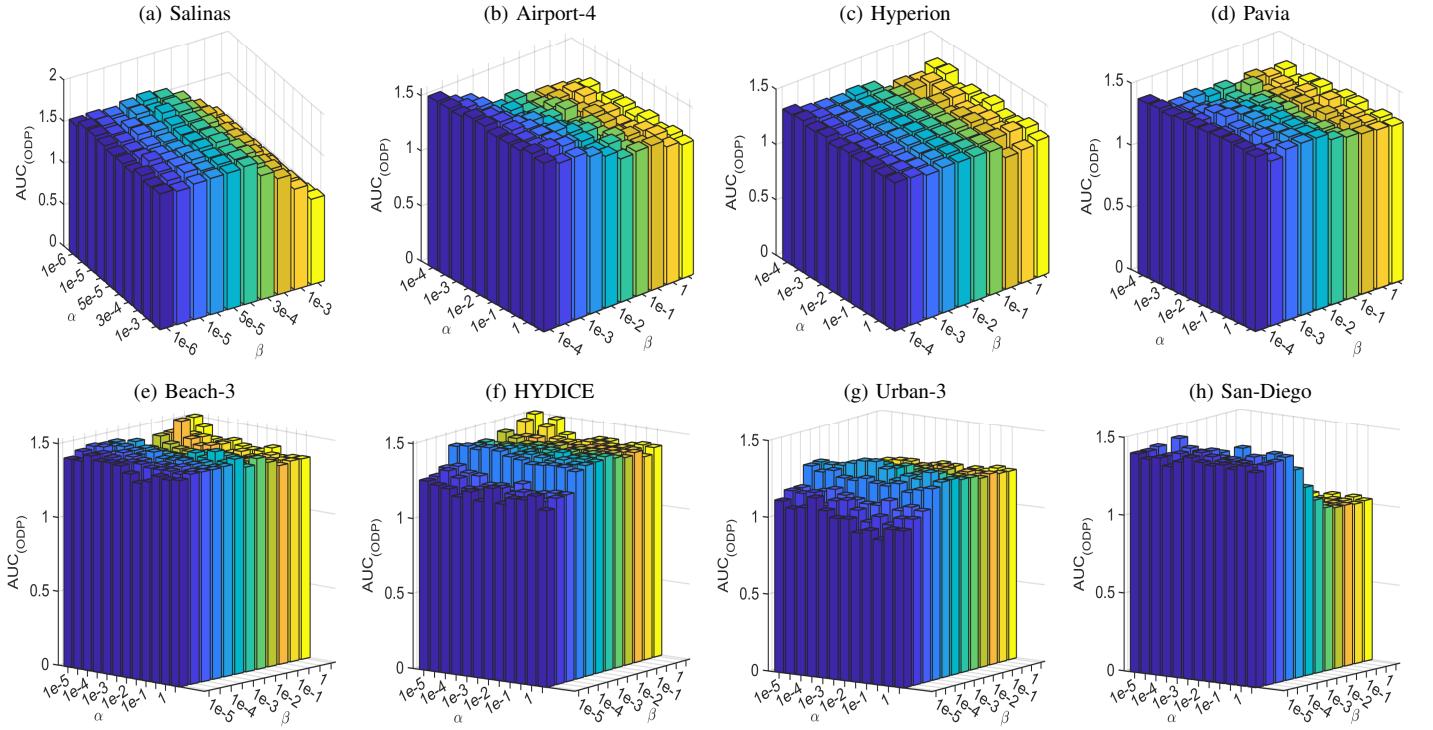


Fig. 5: Sensitivity analysis of the trade-off parameters α and β . For each HSI data, the TR rank is fixed to $(r_1, r_2, r_3) = (6, 6, 6)$, and the nonconvex combination $\Phi+\psi$ is set to be the same, i.e., $\Phi(\cdot) = \psi(\cdot)$ = Lp.

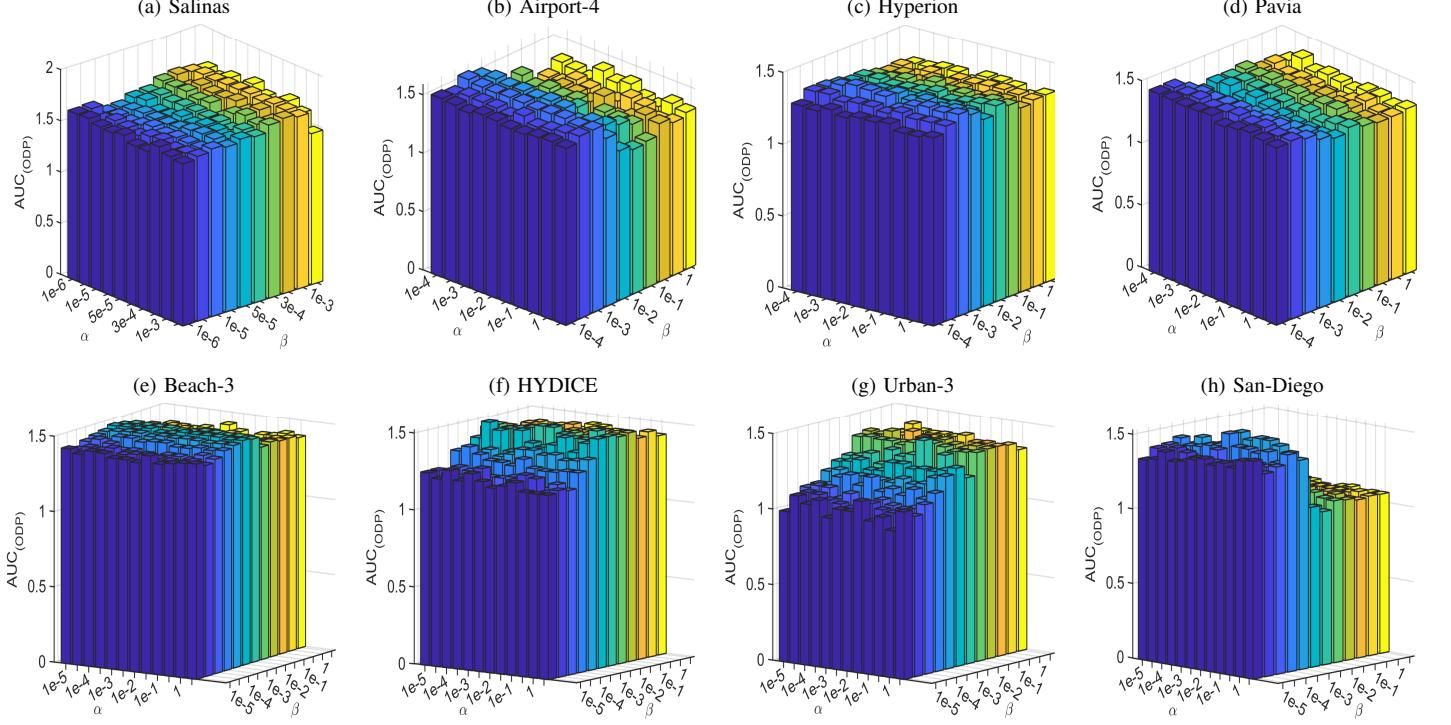


Fig. 6: Sensitivity analysis of the trade-off parameters α and β . For each HSI data, the TR rank is fixed to $(r_1, r_2, r_3) = (6, 6, 6)$, and the nonconvex combination $\Phi + \psi$ is set to be the same, i.e., $\Phi(\cdot) = \psi(\cdot) = \text{Capped-Lp}$.

IV. FEATURE VISUALISATION REGARDING EACH MODULE OF THE PROPOSED METHOD

In this section, we utilize feature visualisation to illustrate the specific effects of each module (i.e., nonconvex regularization schemes, prior representation paradigms, and tensor decomposition schemes) on background suppression and anomaly enhancement.

- Feature visualisation using different nonconvex regularization schemes are displayed in Figures 7-9. These visual effects illustrate that compared with other nonconvex degenerate versions (TCTV, GNTCTV, EGNTCTV, UNTRFR), the proposed regularization method (i.e., EUNTRFR) exhibits better performance in simultaneously suppressing the background and enhancing the detection of anomalous targets.
- Feature visualisation utilizing different prior representation schemes are presented in Figure 10. From these visual results, it can be observed that the HAD methods induced by the three prior representation paradigms are all capable of detecting relatively clear anomalous targets. Whereas, compared to the pure low-rankness method and the pure smoothness method, the proposed joint L+S method can better suppress the background.
- Feature visualisation using different tensor decomposition schemes are provided in Figures 11-12. From these results, we can clearly observe that: among all decomposition strategies, although the MERA decomposition is capable of effectively suppressing background components, the detected anomalous objects are not clearly discernible in most cases. Compared with the remaining decomposition methods (e.g., TT, Tucker, T-SVD), our approach detects most anomalous objects while experiencing less interference from background components.

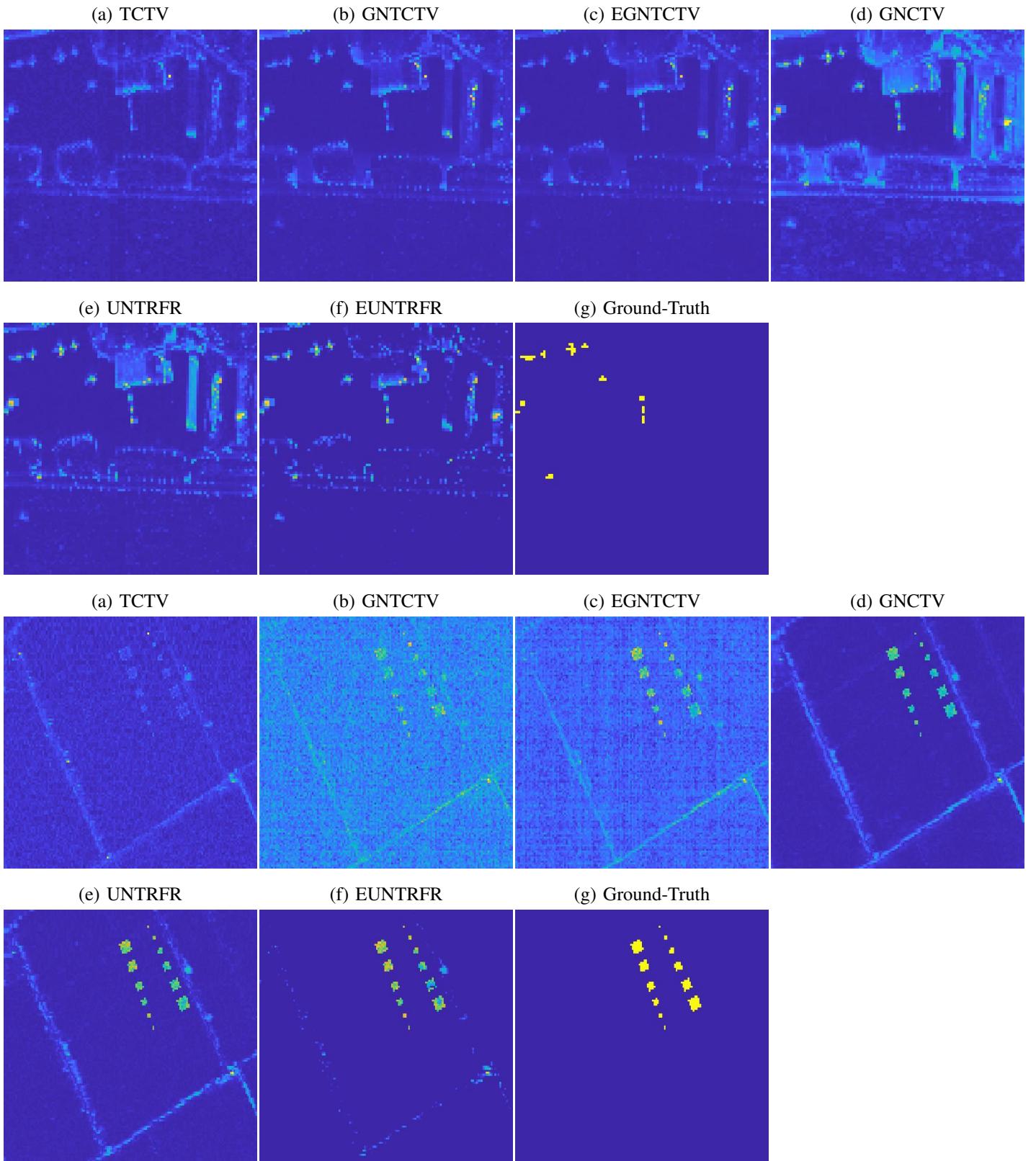


Fig. 7: **Feature Visualisation Using Different Nonconvex Regularization Schemes (Example 1):** The impact of adopting various nonconvex regularization schemes to encode the prior structures of background tensor on the performance of anomaly detection.

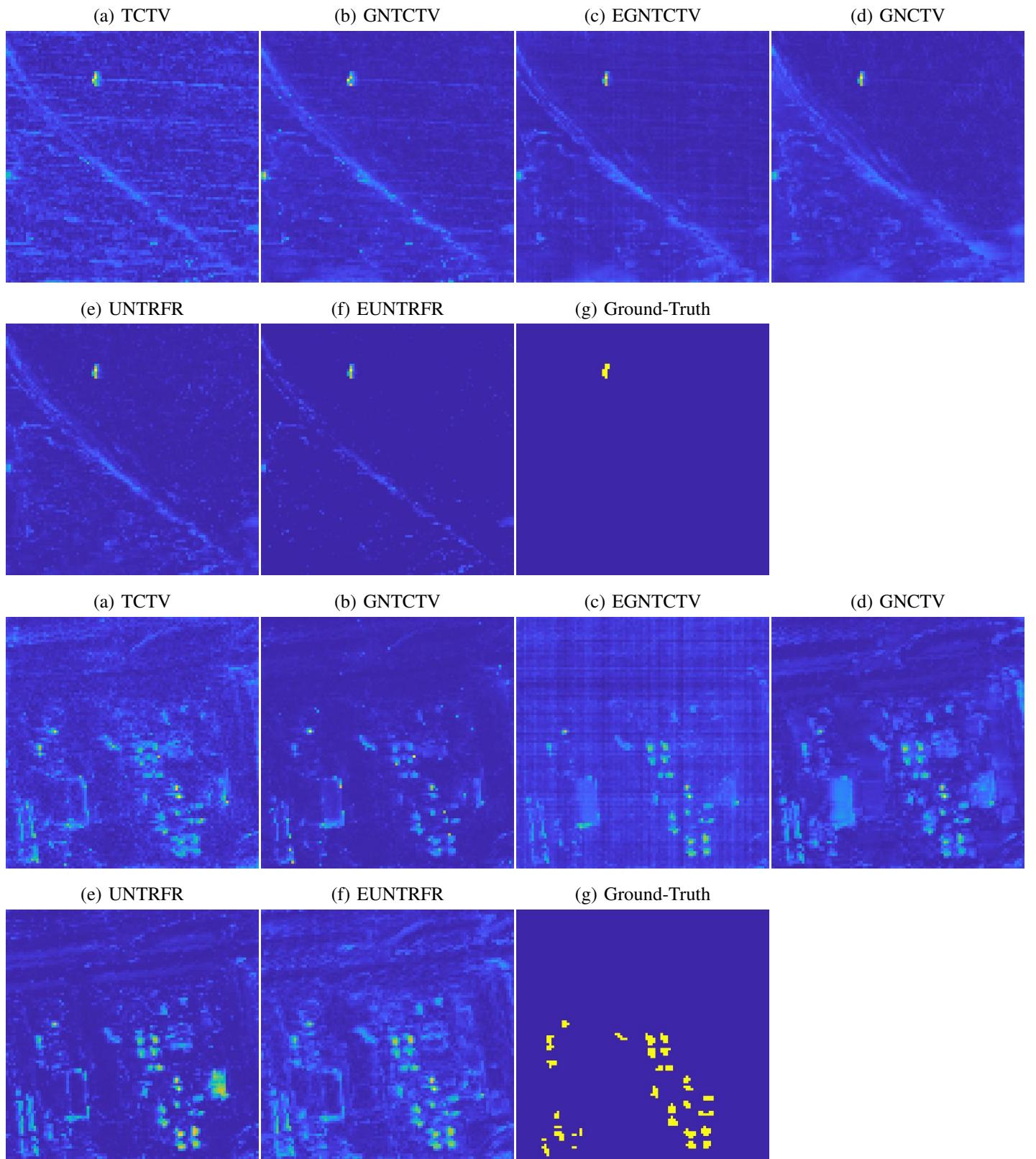


Fig. 8: **Feature Visualisation Using Different Nonconvex Regularization Schemes (Example 2):** The impact of adopting various nonconvex regularization schemes to encode the prior structures of background tensor on the performance of anomaly detection.

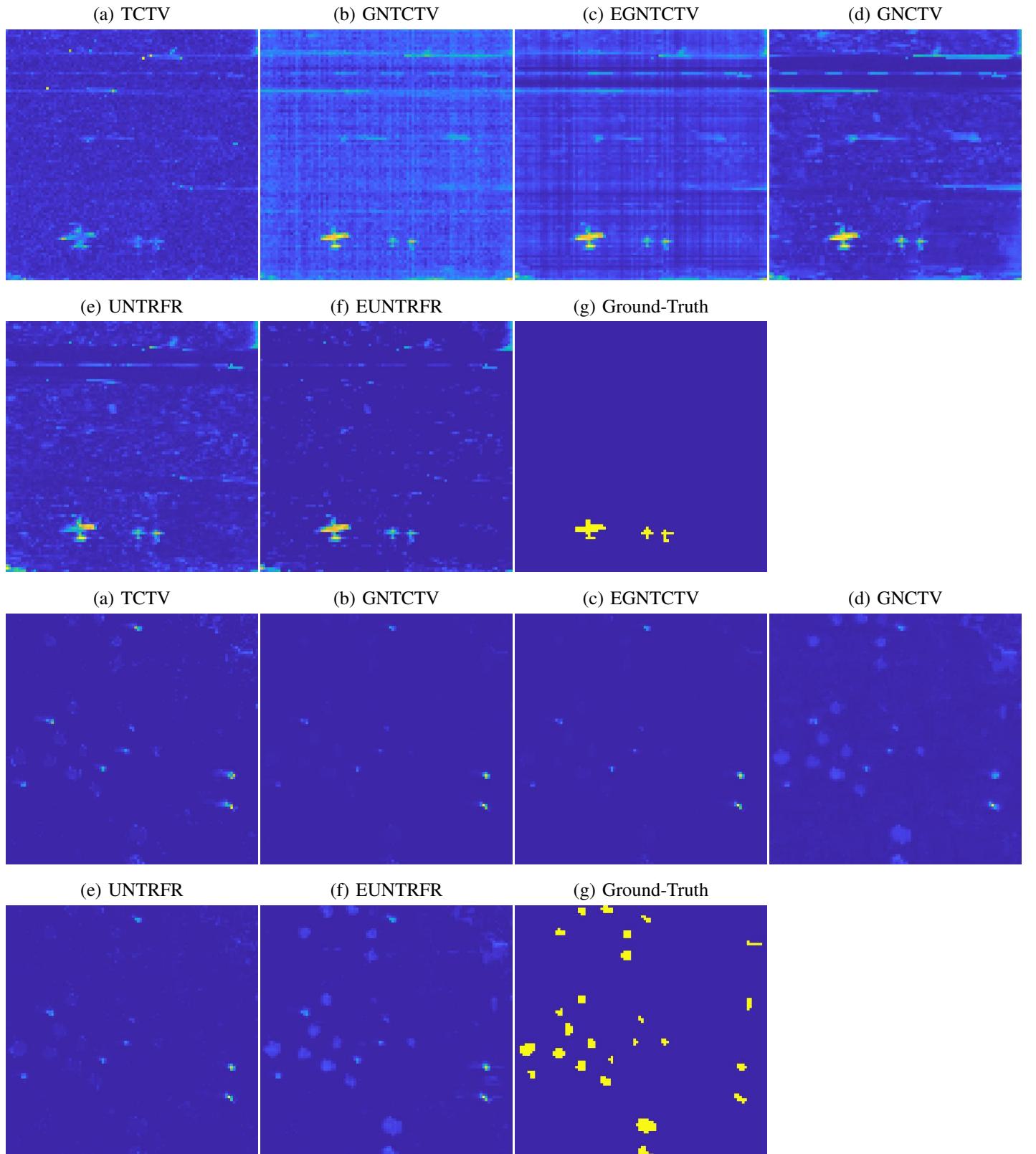


Fig. 9: **Feature Visualisation Using Different Nonconvex Regularization Schemes (Example 3):** The impact of adopting various nonconvex regularization schemes to encode the prior structures of background tensor on the performance of anomaly detection.

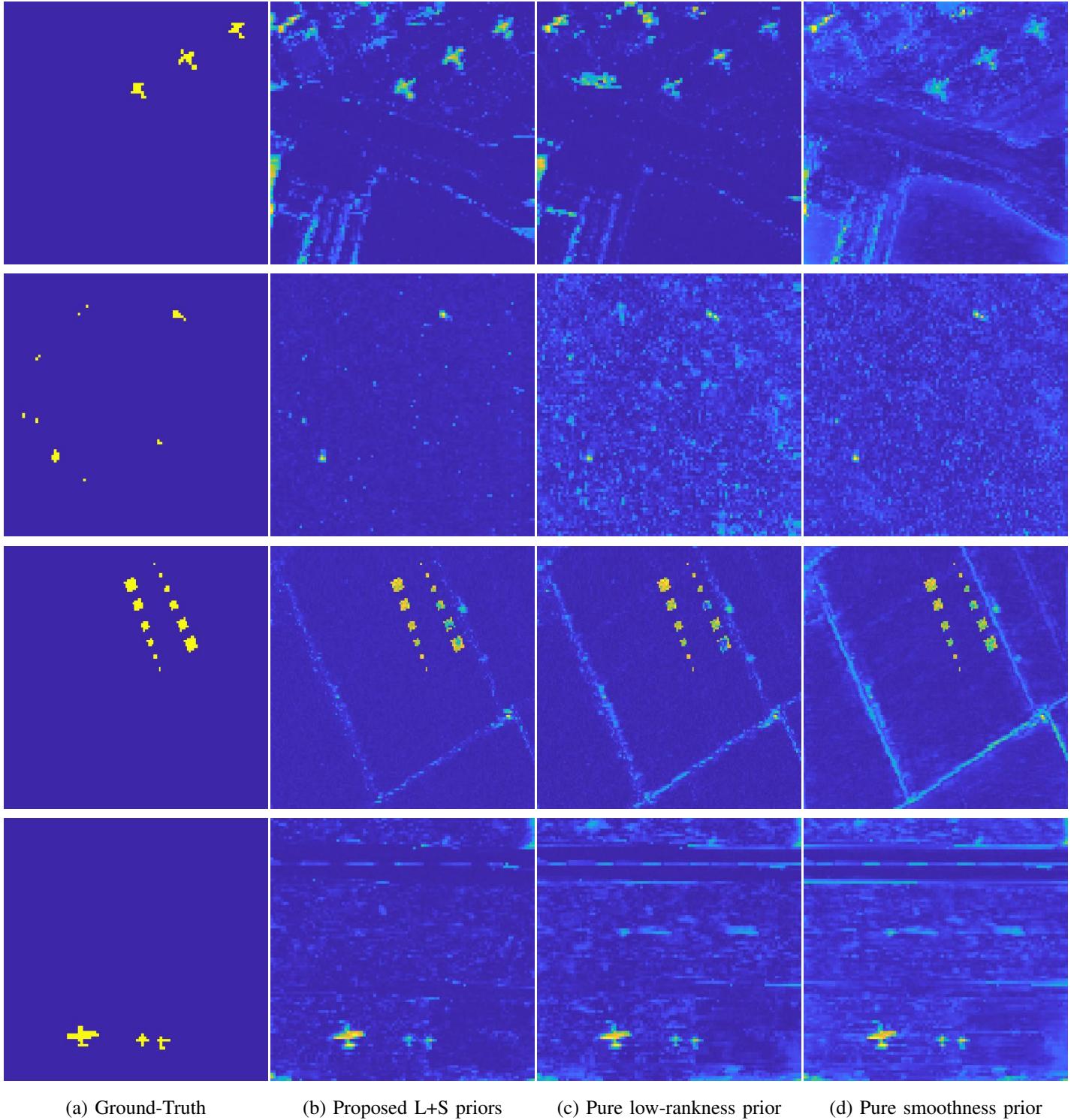


Fig. 10: **Feature Visualisation Utilizing Different Prior Representation Schemes:** The influence of different prior representation methods (i.e., proposed joint low-rankness plus smoothness priors, pure low-rankness prior, pure smoothness prior) upon anomaly detection performance of the proposed method.

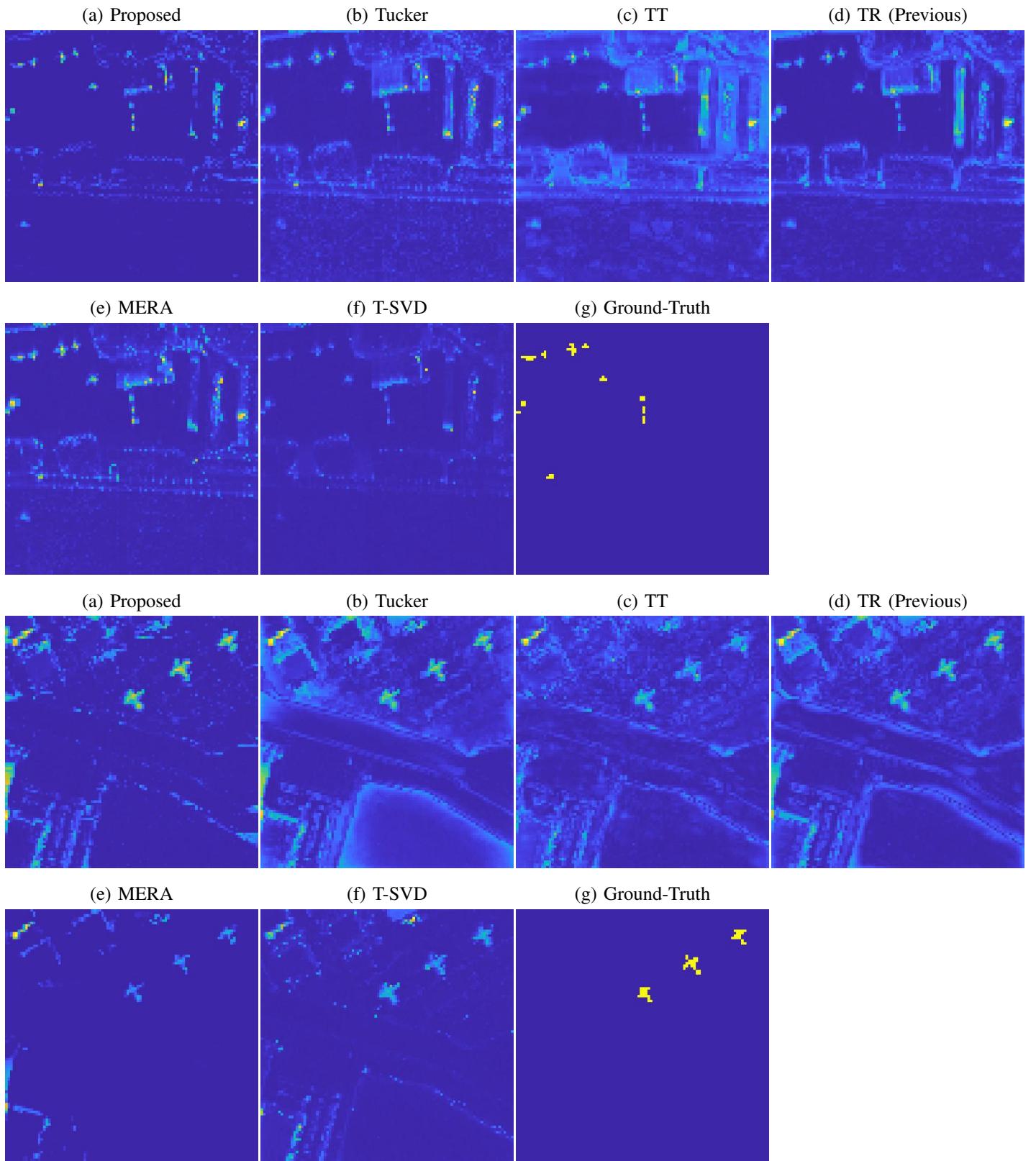


Fig. 11: **Feature Visualisation Using Different Tensor Decomposition Schemes (Example 1):** The influence of various tensor decomposition schemes upon anomaly detection performance of the proposed method.

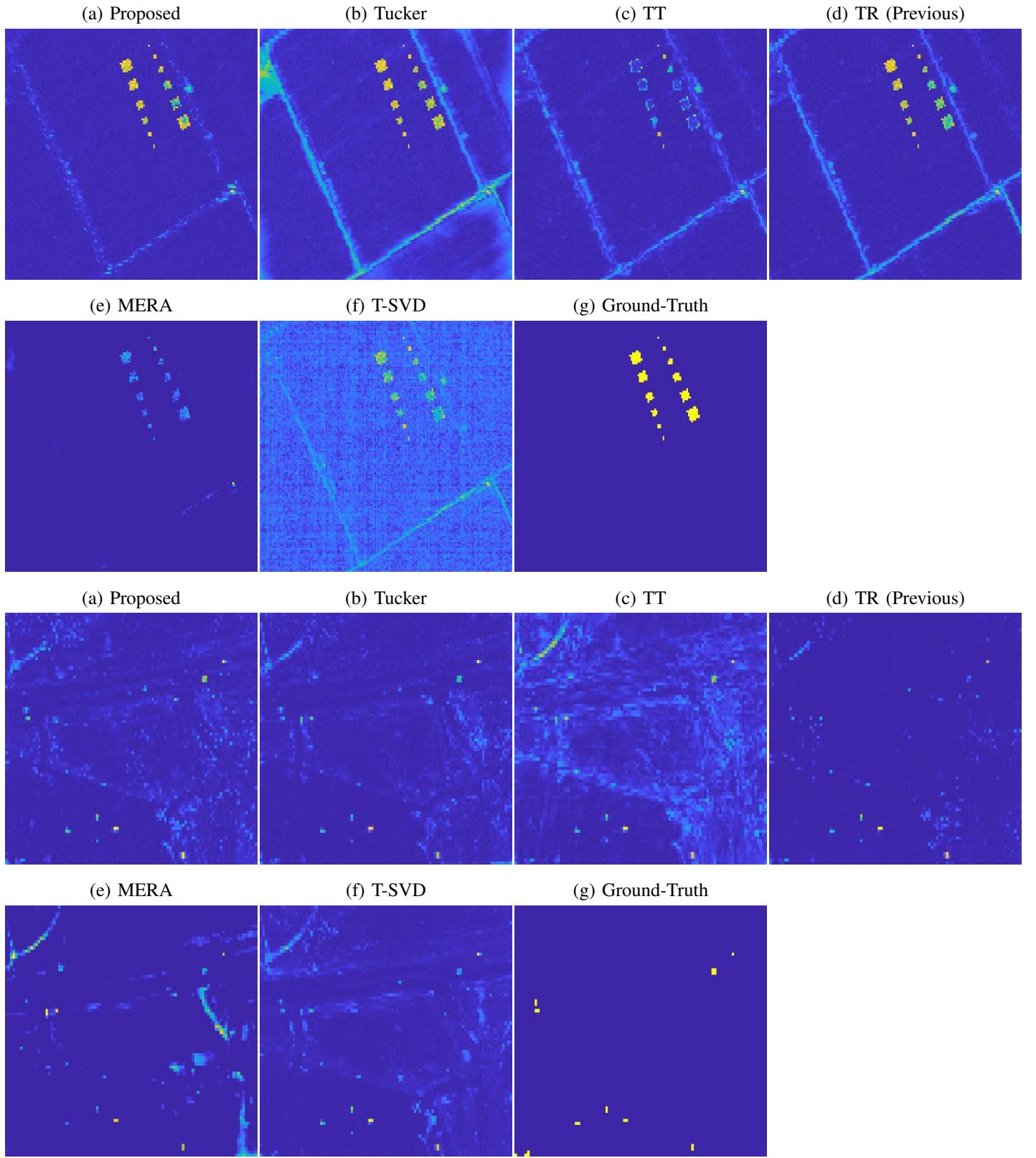


Fig. 12: **Feature Visualisation Using Different Tensor Decomposition Schemes (Example 2):** The influence of various tensor decomposition schemes upon anomaly detection performance of the proposed method.

V. MORE EXPERIMENT RESULTS ON LARGE-SCALE HSI DATASETS

In this section, we mainly present additional quantitative and qualitative results obtained from experiments conducted on the following several large-scale HSI datasets. These these HSI datasets include

- **Qingpu Dataset [16]:** This dataset was acquired over the Shanghai region of China using the Airborne Multi-Modular Imaging Spectrometer (AMMIS). As shown in Figure 13 (a), the Qingpu-I dataset has a size of 600×400 pixels, with four vehicles considered as the anomalies. As shown in Figure 13 (b), the Qingpu-II dataset has a size of 400×740 pixels, with the blue rooftops considered the targets to be detected.
- **Avon Dataset [16]:** This dataset was acquired the ProSpecTIR-VS sensor system, with the acquisition area located in the southern part of Avon, Rochester, New York, USA. As shown in Figure 13 (c), the scene containing 25 grid-patterned tarps and three red or blue felt pads are considered anomalies.
- **Cri dataset [17]:** The Cri dataset was acquired by imaging an area of Wuhan University with the Nuance Cri hyperspectral sensor. The full image contains 400×400 pixels with 46 spectral channels retained at a spectral resolution of 10 nm. Please see Figure 13 (d) for more details.
- **Beach dataset [18]:** This dataset was collected via a drone-mounted hyperspectral line-scan system on a beach in QLD, Australia. It transitions from vegetation, to sand, to shallow ocean, and includes anomalous objects such as different colored and sized tarps, two vehicles, and reflective cones (the smaller yellow objects in the ground truth). Please see Figure 14 for details.
- **SNP dataset [18]:** This dataset was collected using the Sentinel-2 multispectral satellites and downloaded via Sentinel Hub. The dataset covers part of SNP in California and includes live wildfires as anomalies that were identified using infrared bands. Please see Figure 15 (a) for details.
- **Synthetic dataset [18]:** This dataset was created by sampling 7.1-m pixels from an AVIRIS radiance dataset collected near Gulfport, MS, USA. The background is generated from vegetation, sand, and water pixels with smoothed transitions between each type. The anomalies are square targets made from airport runway pixels. Please see Figure 15 (b) for details.

The two-dimensional anomaly detection maps of various HAD methods on the above-mentioned HSI datasets are displayed in Figures 16, 17, 18, 19, 20, 21, 22. Compared with other detectors, the proposed method strikes a superior balance between accurate anomaly recognition and effective background removal. The box-whisker maps of different detectors on several datasets are shown in Figure 23. As can be seen from this figure, the gap between the anomaly box and the background box using the proposed HAD method is either the largest or the second largest. The 3-D ROC curves and 2-D ROC curves of each algorithm on several HSI datasets are presented in Figure 24. As shown in Figure 24 (b), the proposed algorithm outperforms other algorithms in the majority of scenarios, achieving a higher detection rate across various false alarm rates. Taking into account all the aforementioned evaluation metrics, including the anomaly detection map, 3-D ROC curve, AUC values, and box-whisker plot, our proposed HAD algorithm demonstrates good performance.

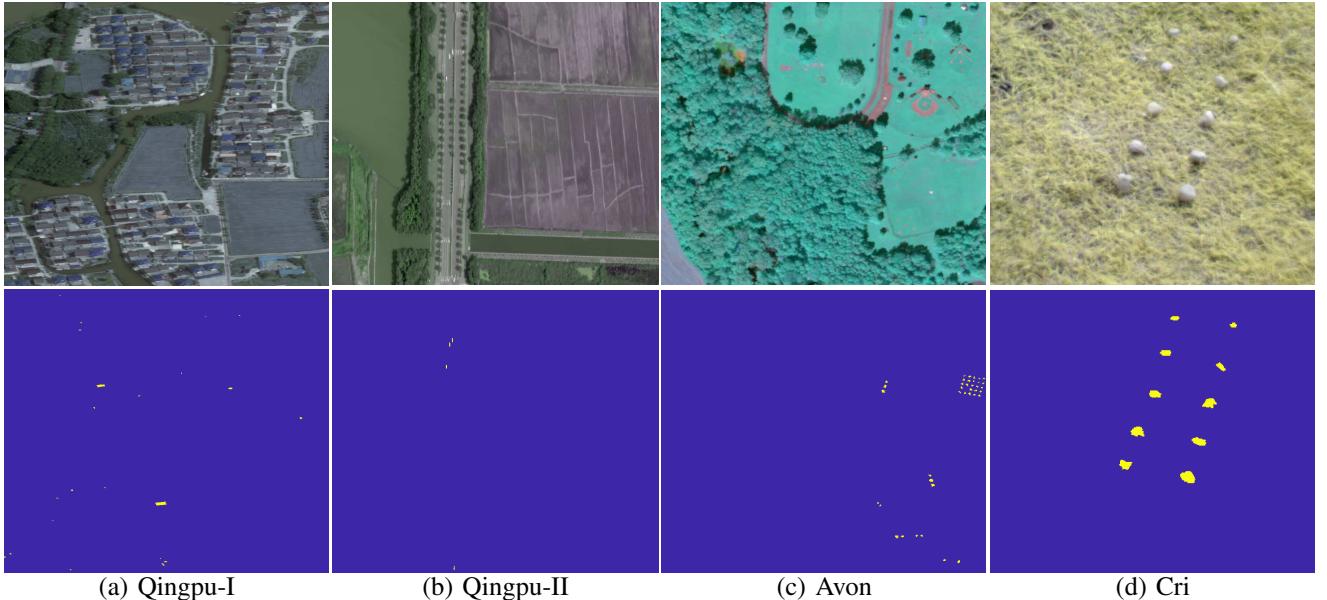


Fig. 13: Pseudocolor images and ground-truth maps of Qingpu-I dataset ($740 \times 400 \times 250$), Qingpu-II dataset ($400 \times 600 \times 250$), Avon dataset ($400 \times 400 \times 360$), and Cri dataset ($400 \times 400 \times 46$). Top: raw dataset RGB image, bottom: anomaly ground-truth image.

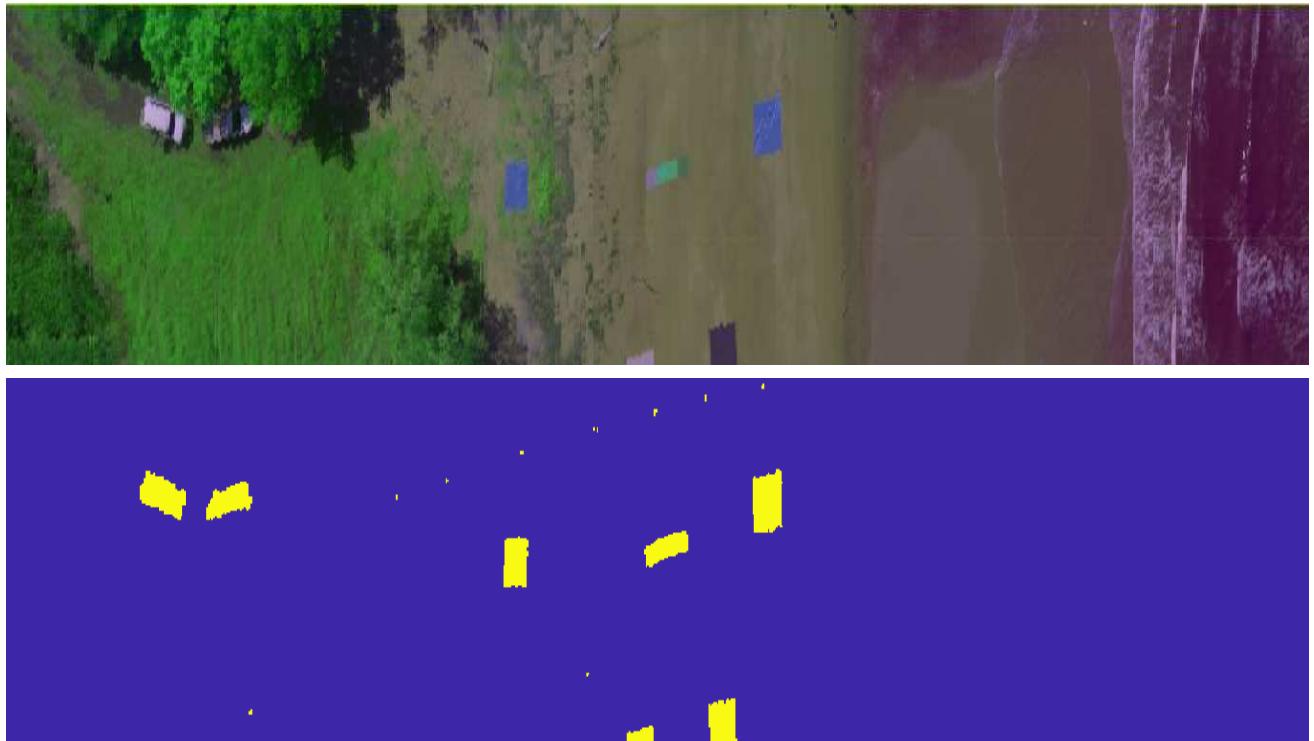


Fig. 14: Pseudocolor images and ground-truth maps of Beach ($452 \times 3072 \times 108$). Top: Beach dataset RGB image, bottom: anomaly ground-truth image.

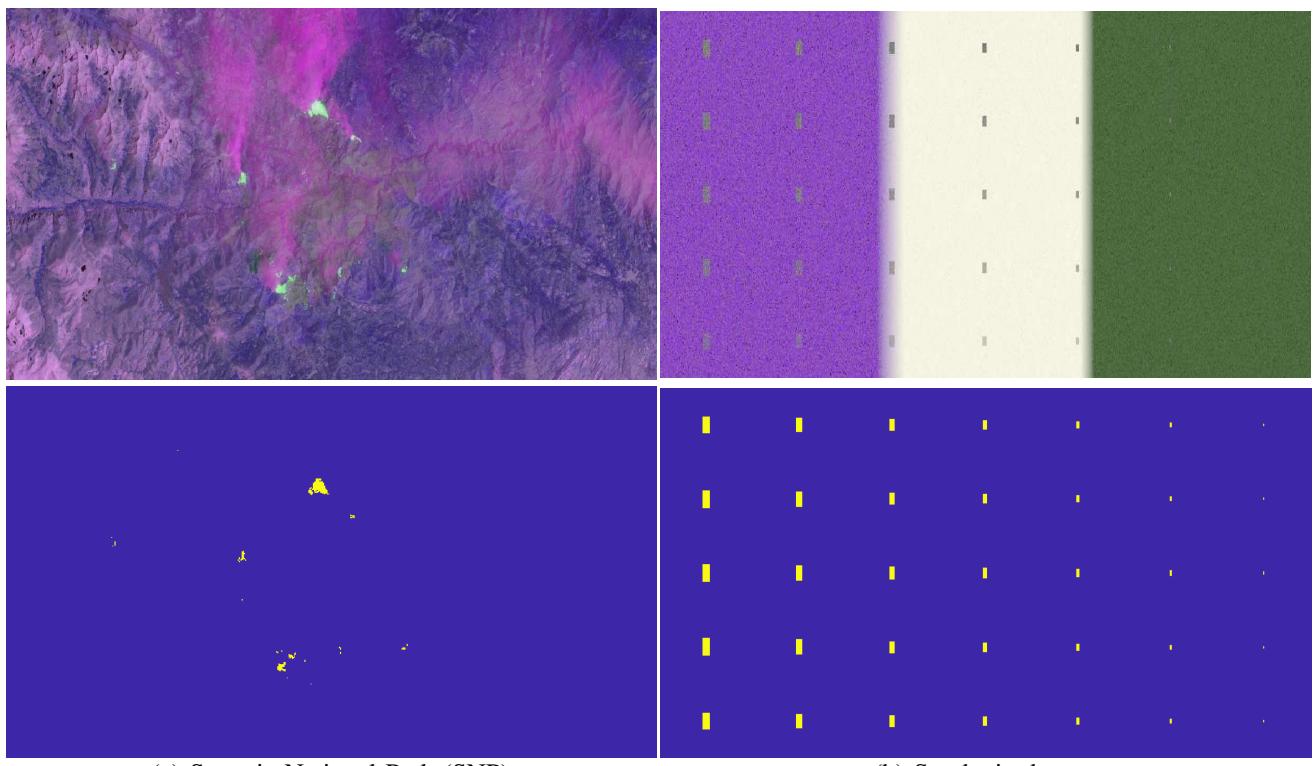


Fig. 15: Pseudocolor images and ground-truth maps of National Park (SNP) ($1116 \times 2499 \times 13$) and Synthetic dataset ($600 \times 2400 \times 90$). Top: raw dataset RGB image, bottom: anomaly ground-truth image.

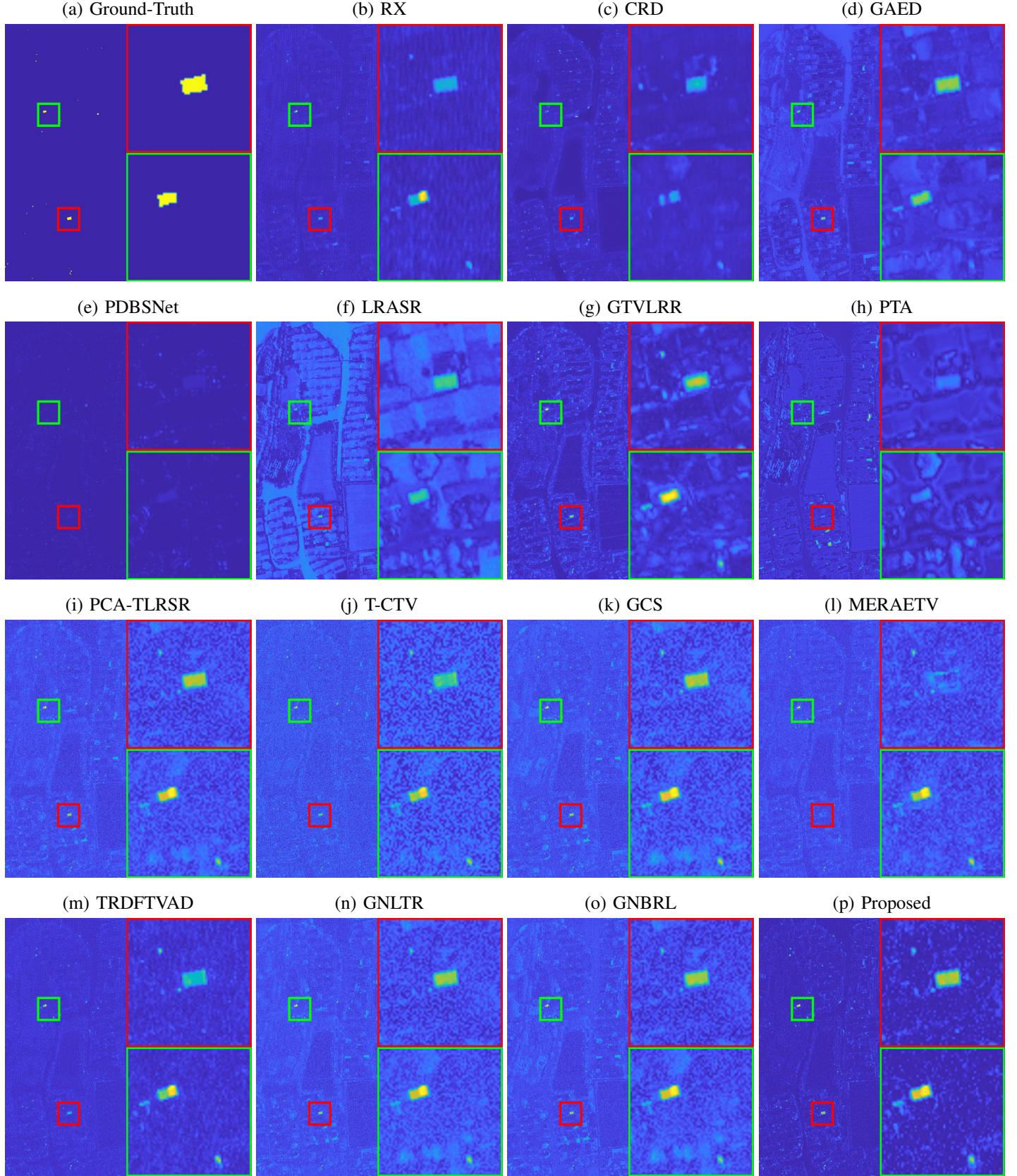


Fig. 16: Anomaly detection map of various HAD methods on HSI dataset named Qingpu-I ($740 \times 400 \times 250$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

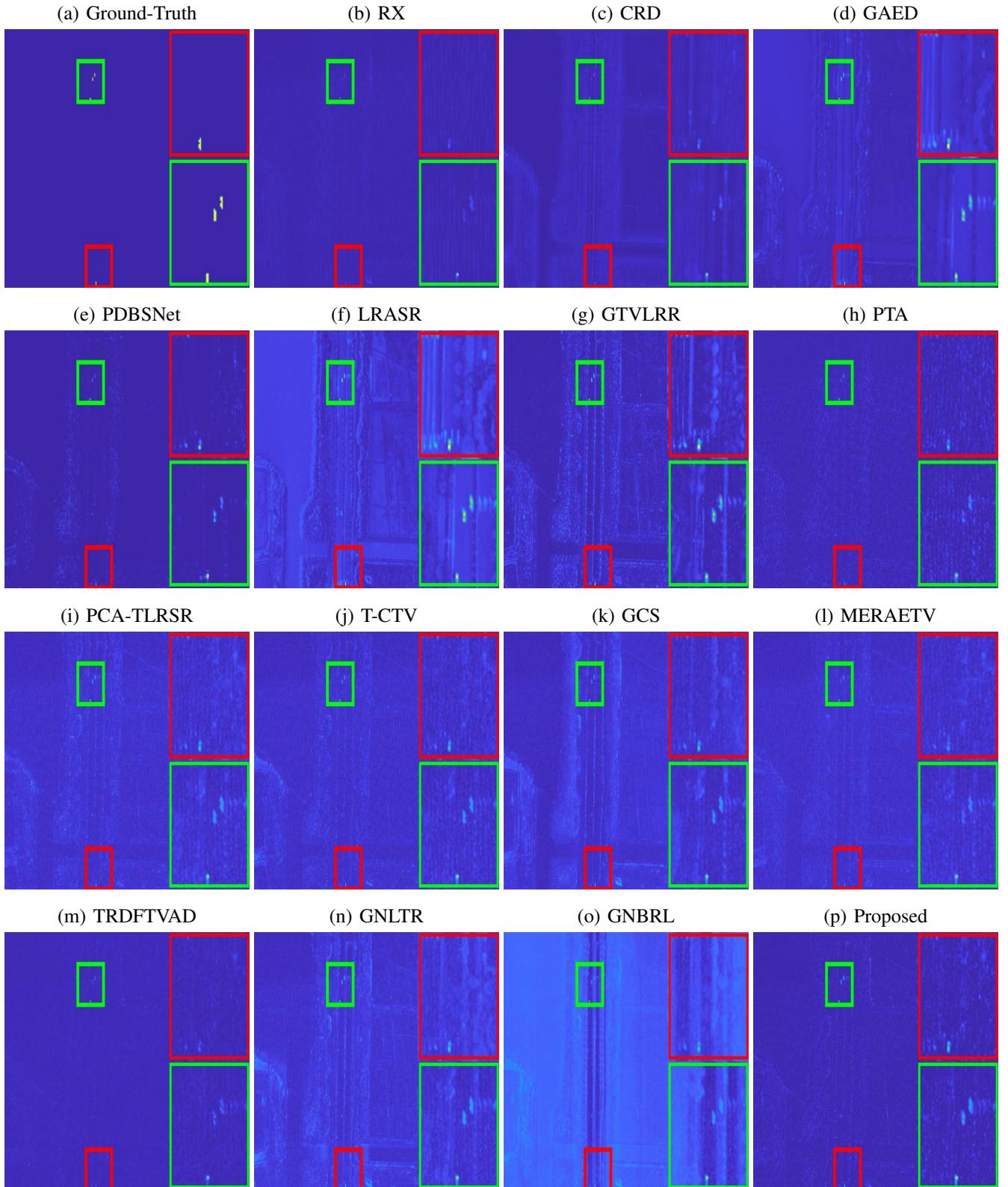


Fig. 17: Anomaly detection map of various HAD methods on HSI dataset named Qingpu-II ($400 \times 600 \times 250$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

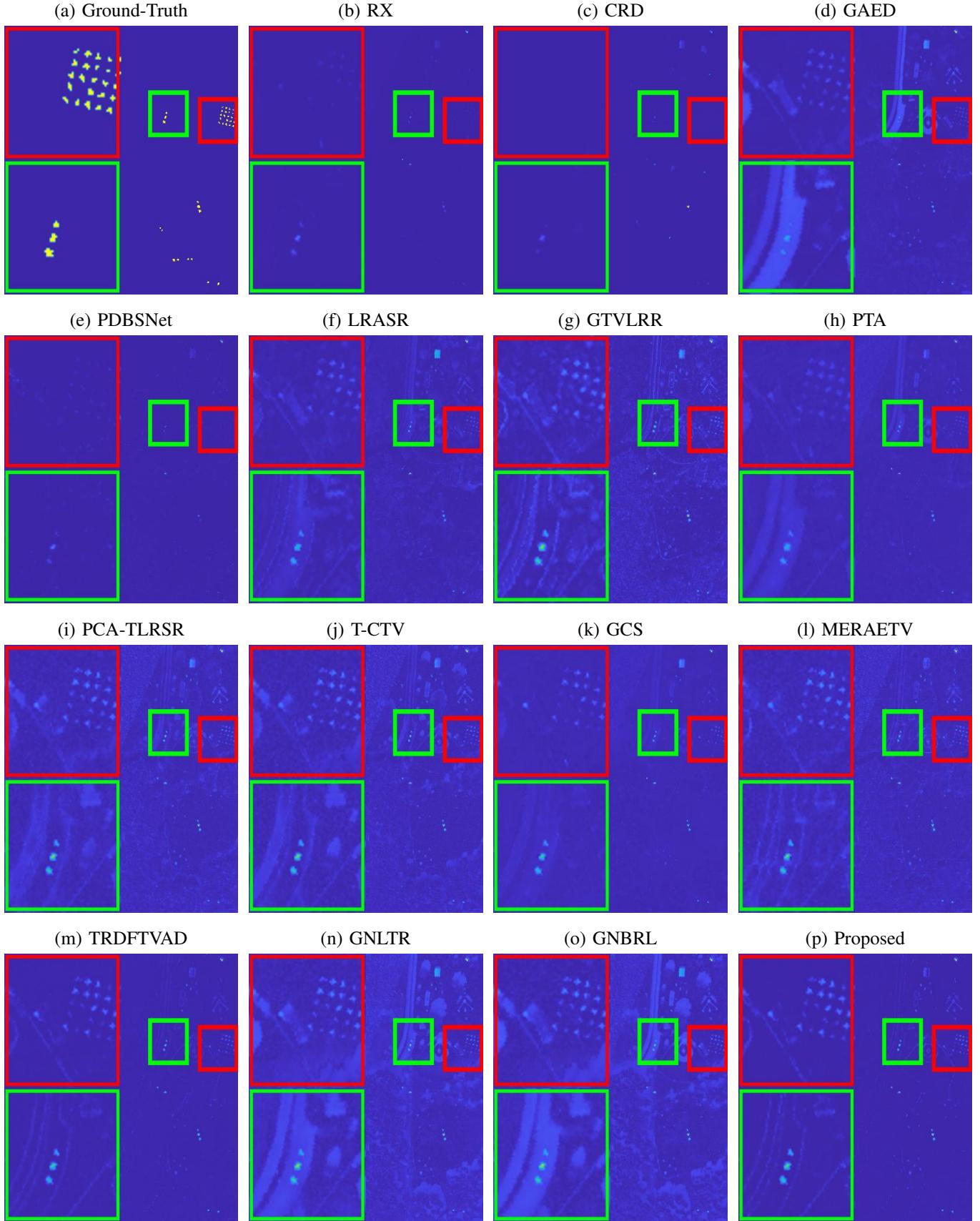


Fig. 18: Anomaly detection map of various HAD methods on large-scale HSI dataset named AVON ($400 \times 400 \times 360$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

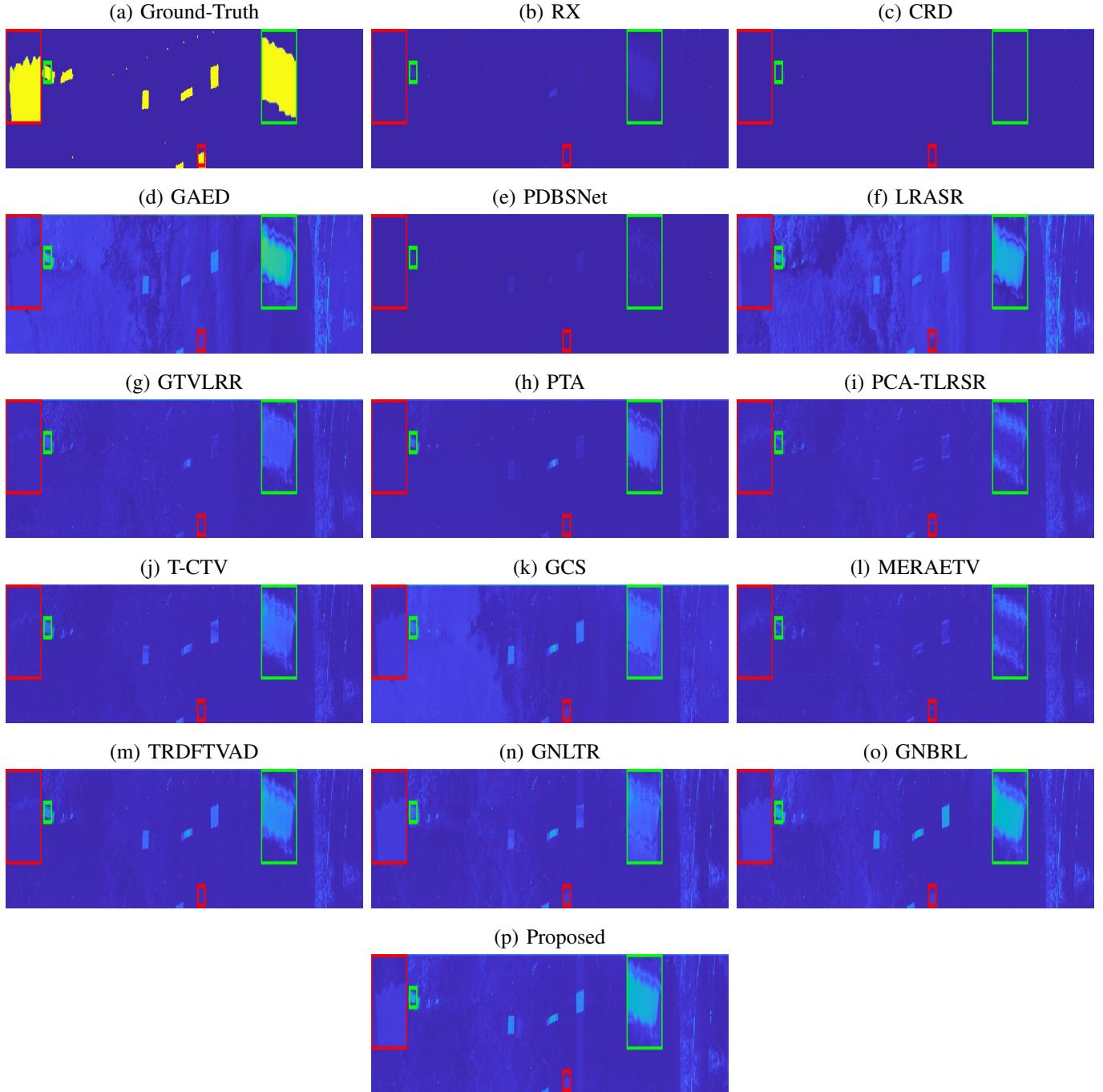


Fig. 19: Anomaly detection map of various HAD methods on HSI dataset named Beach ($452 \times 3072 \times 108$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV. (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

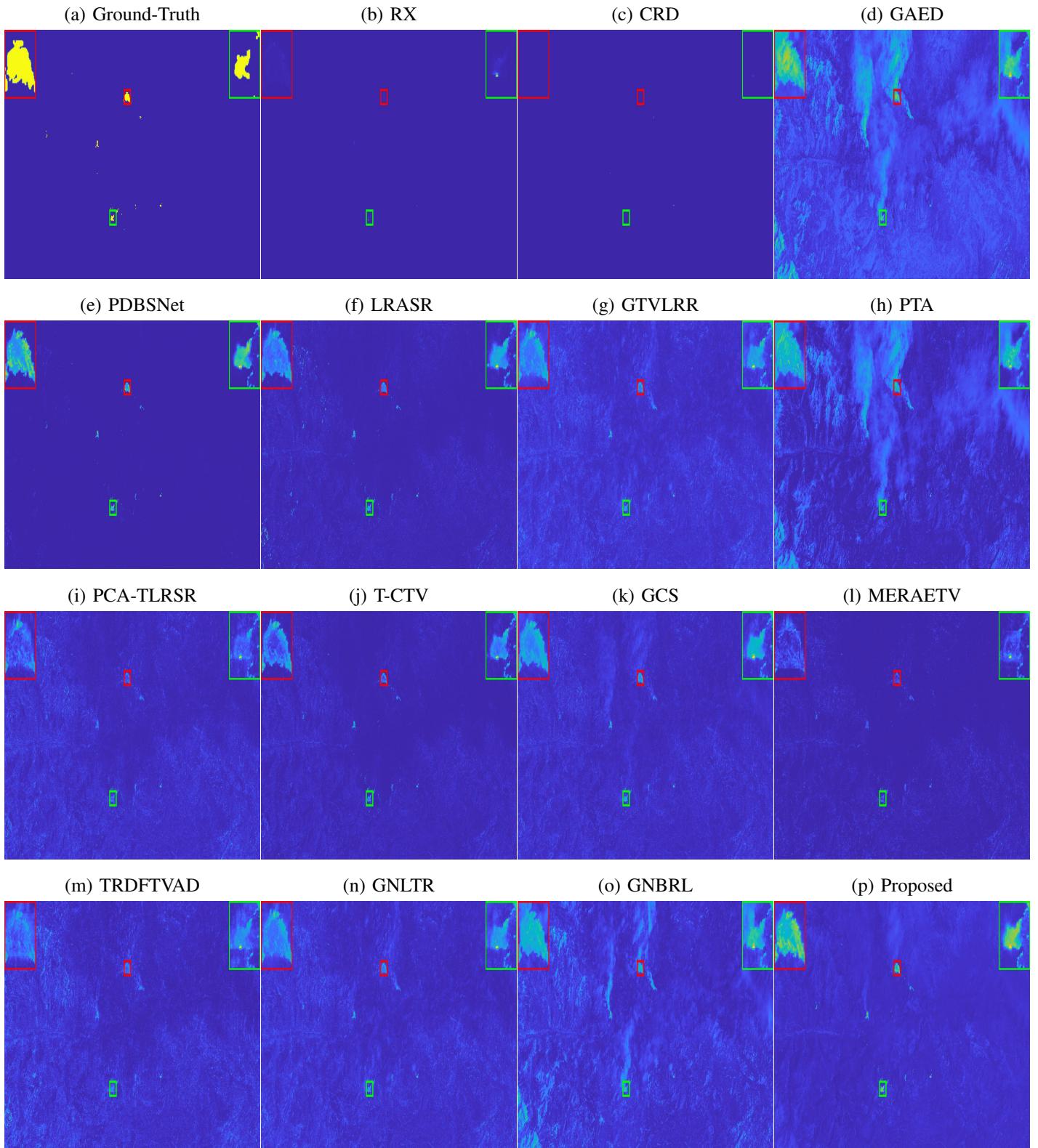


Fig. 20: Anomaly detection map of various HAD methods on large-scale Sequoia National Park (SNP) dataset ($1116 \times 2499 \times 13$).
(a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV.
(k) GCS. (l) MERAETV. (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

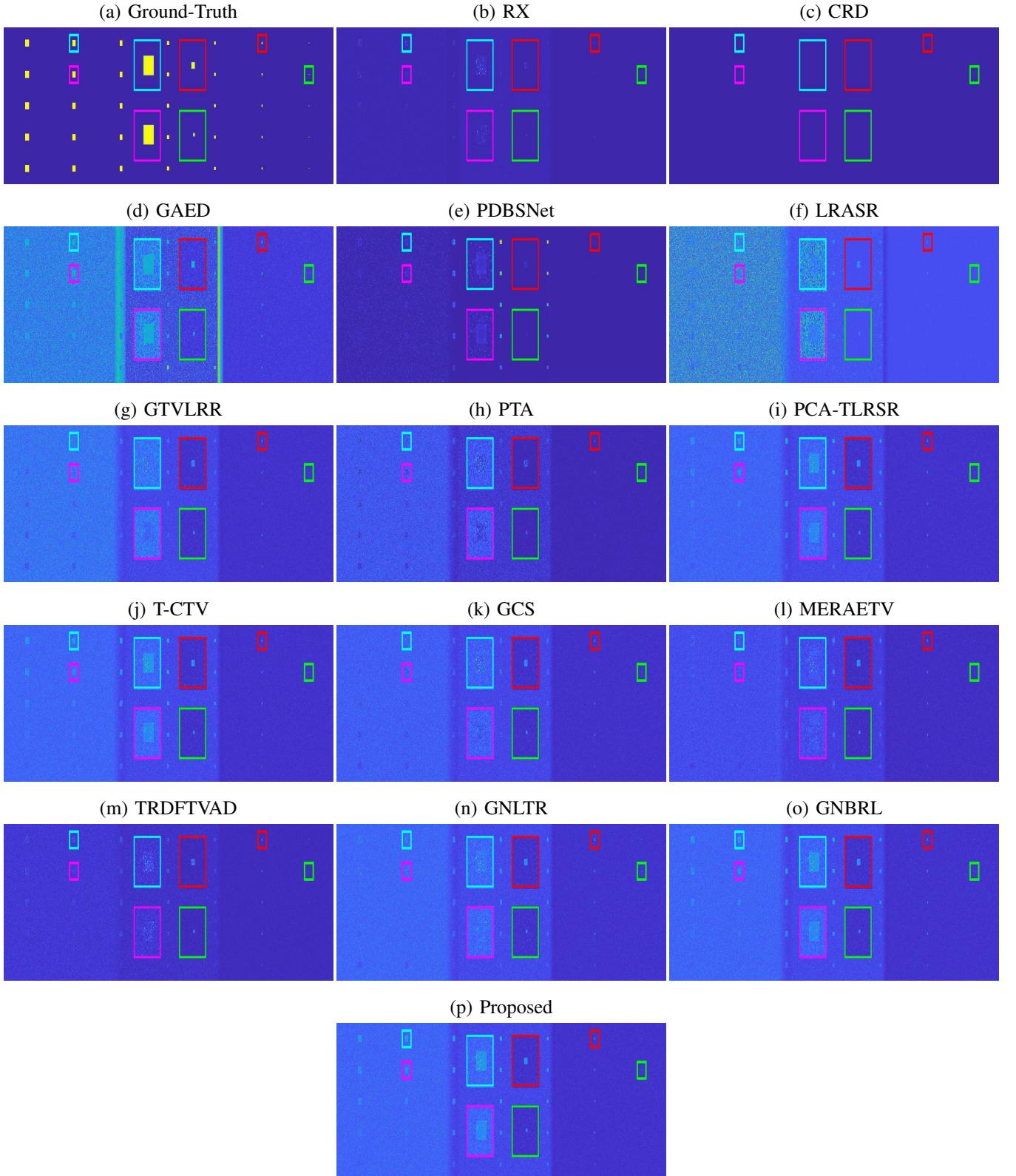


Fig. 21: Anomaly detection map of various HAD methods on synthetic HSI dataset ($600 \times 2400 \times 90$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV. (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

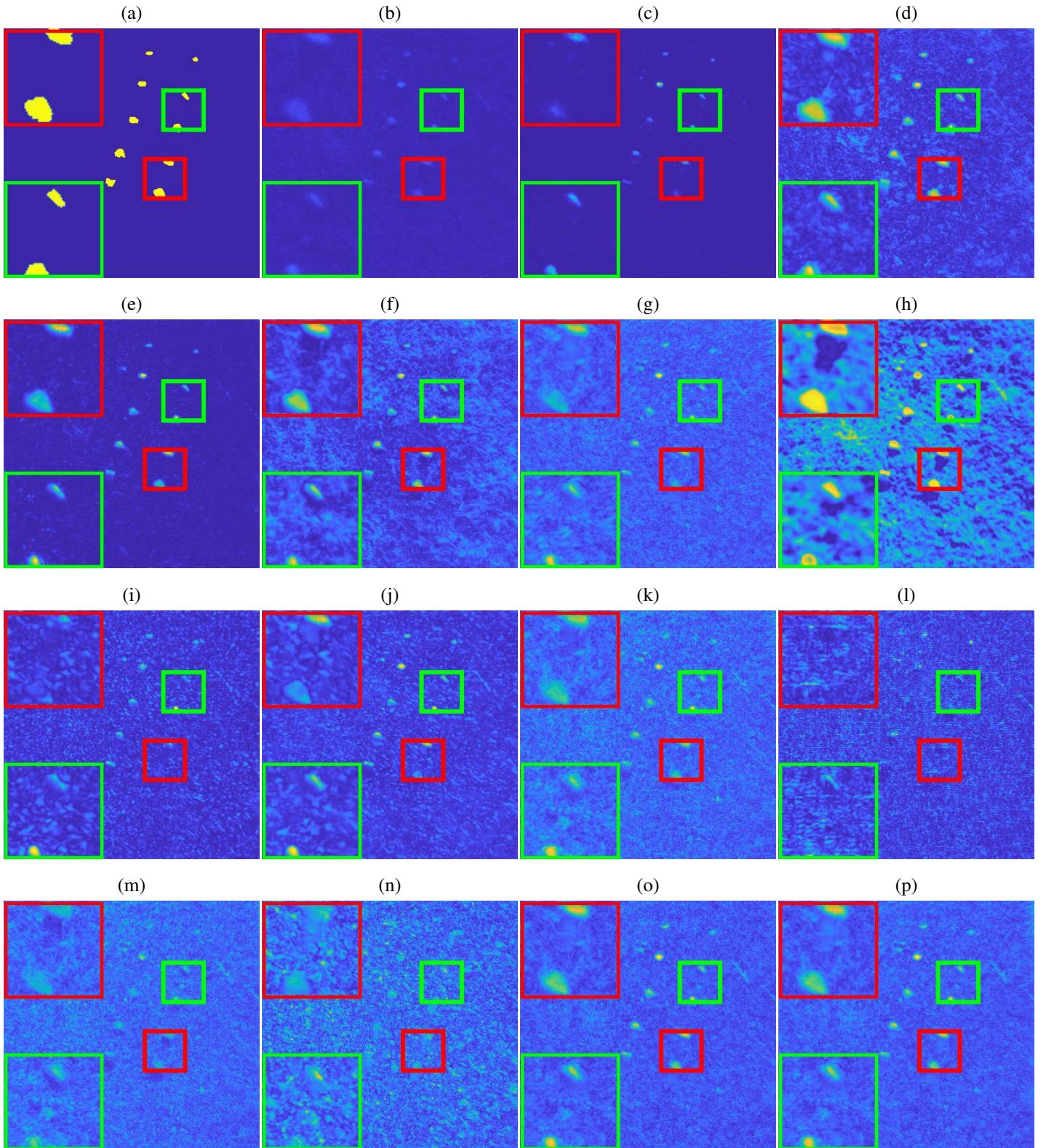


Fig. 22: Anomaly detection map of various HAD methods on HSI dataset named CRI ($400 \times 400 \times 46$). (a) Ground-Truth. (b) RX. (c) CRD. (d) GAED. (e) PDBSNet. (f) LRASR. (g) GTVLRR. (h) PTA. (i) PCA-TLRSR. (j) T-CTV. (k) GCS. (l) MERAETV. (m) TRDFTVAD. (n) GNLTR. (o) GNBRL. (p) Proposed.

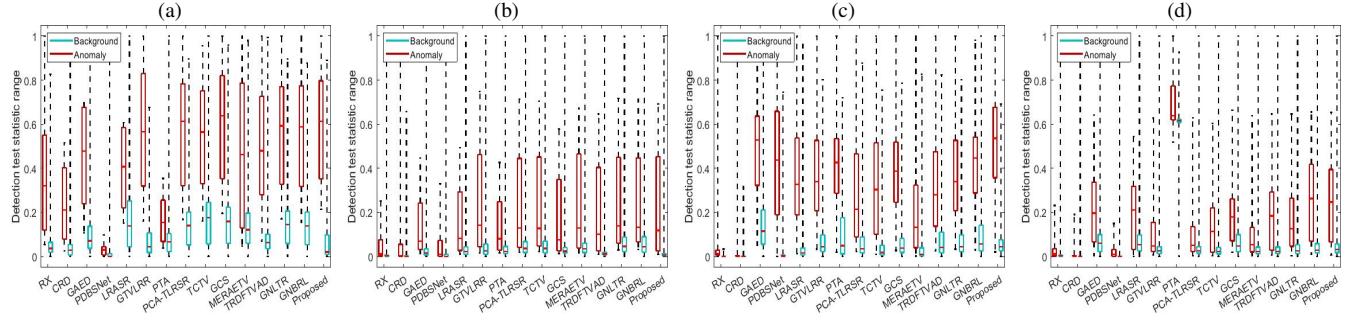


Fig. 23: Separability maps of various HAD methods for different HSI datasets. (a) Qingpu-I. (b) Avon. (c) SNP. (d) Beach.

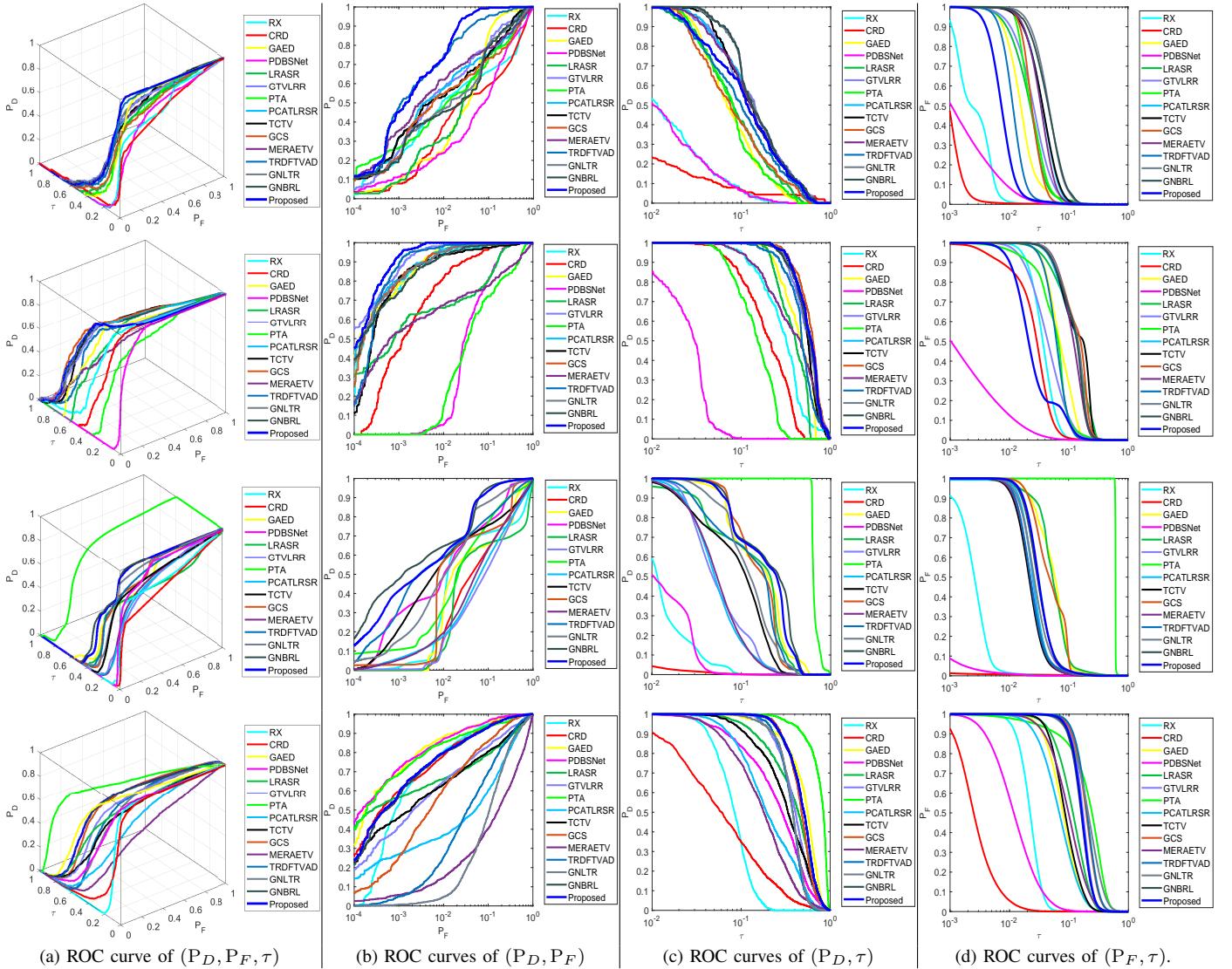


Fig. 24: 3-D and 2-D ROC curves' performance comparison of different HAD approaches on large-scale HSI datasets: Avon, Qingpu-I, Beach, CRI (from top to bottom).

REFERENCES

- [1] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, 1990.
- [2] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463–1474, 2014.
- [3] P. Xiang, S. Ali, S. K. Jung, and H. Zhou, "Hyperspectral anomaly detection with guided autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5538818.
- [4] Y. Zhang, B. Du, L. Zhang, and S. Wang, "A low-rank and sparse matrix decomposition-based mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1376–1389, 2015.
- [5] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, 2015.
- [6] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, 2019.
- [7] L. Li, W. Li, Y. Qu, C. Zhao, R. Tao, and Q. Du, "Prior-based tensor approximation for anomaly detection in hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1037–1050, 2020.
- [8] M. Wang, Q. Wang, D. Hong, S. K. Roy, and J. Chanussot, "Learning tensor low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 679–691, 2022.
- [9] H. Wang, J. Peng, W. Qin, J. Wang, and D. Meng, "Guaranteed tensor recovery fused low-rankness and smoothness," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10990–11007, 2023.
- [10] M. Feng, W. Chen, Y. Yang, Q. Shu, H. Li, and Y. Huang, "Hyperspectral anomaly detection based on tensor ring decomposition with factors TV regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5514114.
- [11] H. Qin, Q. Shen, H. Zeng, Y. Chen, and G. Lu, "Generalized nonconvex low-rank tensor representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5526612.
- [12] Q. Yu and M. Bai, "Generalized nonconvex hyperspectral anomaly detection via background representation learning with dictionary constraint," *SIAM J. Imag. Sci.*, vol. 17, no. 2, pp. 917–950, 2024.
- [13] Q. Xiao, L. Zhao, S. Chen, and X. Li, "Hyperspectral anomaly detection via MERA decomposition and enhanced total variation regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5514919.
- [14] D. Wang, L. Zhuang, L. Gao, X. Sun, M. Huang, and A. Plaza, "PDBSNet: pixel-shuffle down-sampling blind-spot reconstruction network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2023, Art. no. 5511914.
- [15] J. Lian, L. Wang, H. Sun, and H. Huang, "GT-HAD: gated transformer for hyperspectral anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 2, pp. 3631–3645, 2025.
- [16] Z. Wang, K. Tan, X. Wang, and W. Zhang, "Leveraging multi-class background description and token dictionary representation for hyperspectral anomaly detection," *Pattern Recognit.*, p. 111945, 2025.
- [17] D. Zhu, B. Du, and L. Zhang, "Two-stream convolutional networks for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6907–6921, 2020.
- [18] S. Garske, B. Evans, C. Artlett, and K. C. Wong, "Erx: A fast real-time anomaly detection algorithm for hyperspectral line scanning," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5503617.