Journal of Systems Engineering and Electronics
Vol. 37, No. 3, June 2026, pp.779 – 787

Structured robust principal component analysis for infrared small target detection

ZHANG Yongqiang[1,2], LI Yongji[3], CAI Meng[2], ZHANG Ye[1,4], and TAN Yong[1,4,*]

1. School of Physics, Changchun University of Science and Technology, Changchun 130022, China;
2. Luoyang Institute of Electro-optical Equipment, AVIC, Luoyang 471000, China;
3. School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China;
4. Jilin Key Laboratory of Spectral Detection Science and Technology, Changchun 130022, China

Abstract: Extracting infrared small targets from heterogeneous backgrounds remains a challenging task, as these targets lack salient texture and morphological features while the backgrounds are cluttered with noise. Therefore, effectively extracting discriminative features is essential for achieving complete and accurate detection. To address this issue, this paper proposes an algorithmic framework based on robust principal component analysis (RPCA), specifically designed for infrared small target detection in complex backgrounds. First, infrared small target detection is formulated as a generalized RPCA problem. A discriminative and reconstructive dictionary is constructed using supervised learning. Next, by introducing an ideal regularization term, the infrared image is reconstructed without losing structural information, yielding a discriminative principal component representation with respect to the learned dictionary. Finally, the structural features of infrared small targets are reconstructed via sparse coding, thereby enabling the extraction of infrared small targets. Extensive experimental results demonstrate the effectiveness of the proposed method.

Keywords: infrared small target detection, heterogeneous background, robust principal component analysis, sparse coding.

DOI: 10.23919/JSEE.2026.000093

1. Introduction

Infrared small target detection has become a key research focus in computer vision, aiming to automatically localize and identify objects of interest within visual data. It serves as a critical component in a wide range of intelligent systems. This technology has seen extensive practical applications, including the automated recognition and tracking of suspicious individuals in intelligent surveillance systems, as well as the perception of vehicles and pedestrians in autonomous driving scenarios to enhance operational safety. Owing to its capability to effectively model complex visual scenes, object detection plays an indispensable role in advancing industrial applications and cutting-edge intelligent technologies. Despite the inherent advantages of infrared imaging in challenging environments, achieving reliable and real-time detection of small infrared targets remains a significant challenge in practice. On the one hand, such targets are typically characterized by limited spatial extent and extremely weak texture information, resulting in a scarcity of discriminative geometric and structural cues, which complicates robust identification. On the other hand, infrared imagery is often affected by low signal-to-noise ratios and pronounced background noise, further diminishing the contrast between targets and their surrounding regions. Moreover, complex natural backgrounds, such as cloud boundaries and sea clutter, tend to introduce strong interference, substantially increasing false alarm rates and ultimately constraining the overall performance of detection algorithms.

Single-frame infrared small-target detection methods are typically divided into two main categories: local contrast-based filtering approaches and global property-driven convex optimization models. Early single-frame techniques, including the top-hat operator, maximum mean or median filtering strategies, and two-dimensional minimum mean square error adaptive filters, are mainly designed under the assumption that background regions are locally smooth and homogeneous. Based on this assumption, background components are modeled and estimated within local neighborhoods to suppress clutter and highlight potential targets [1,2]. Following the same background consistency principle, local contrast operators [3] and multi-scale contrast-based methods [4] fur-

Manuscript received February 02, 2026.
*Corresponding author.
This work was supported by Jilin Provincial Science and Technology Development Program Project (20260201028GX).

ther enhance target saliency by amplifying the intensity discrepancy between small targets and their surrounding regions.

With the rapid advancement of artificial intelligence, deep learning has triggered a paradigm shift in infrared small target detection, gradually steering research from traditional model-driven frameworks toward data-driven learning-based approaches [5]. Leveraging the powerful representation learning capability of convolutional neural networks (CNNs), these models are able to automatically extract high-level semantic features from large-scale infrared datasets, leading to notable improvements in detection accuracy. Some studies reformulate infrared small target detection as an image segmentation problem [6], where pixel-wise prediction enables precise localization of target regions. However, segmentation-based approaches typically require dense inference over the entire image, involving extensive pixel-level computations, which results in high computational cost and stringent hardware requirements, thereby limiting their applicability in real-time scenarios. Moreover, the emphasis on per-pixel classification often leads to insufficient modeling of global target structure and semantic context, which can degrade performance in target-level representation and category discrimination. Alternatively, infrared small target detection has also been directly formulated as an object detection task [7], adopting well-established deep learning-based detection frameworks. Representative examples include faster regional CNN [8] and the "you only look one" (YOLO) family of models [9], which have been extensively explored in this domain. These approaches exploit hierarchical convolutional architectures with successive down-sampling operations to jointly capture fine-grained local details and global semantic information in infrared imagery. As a result, target localization and category prediction can be performed within a unified detection framework, enabling automated infrared small target detection.

In recent years, graph neural networks (GNNs), which exhibit strong capability in modeling local structural information, have attracted increasing attention in infrared small target detection. Chen et al. [10] constructed a graph representation with super-pixels as nodes by integrating shallow and deep spectral features, and developed a GNN framework that alternates between graph convolution operations and attention mechanisms to simultaneously capture local neighborhood interactions and long-range dependencies. From another perspective, the adaptive residual contrast deep (ARCD) network [11] employs dynamic graph modeling and multi-scale graph convolution to effectively represent fine-grained local features. Despite the significant progress of neural network-based approaches, their dependence on end-to-end feature embedding inevitably results in partial information degradation during feature propagation. This limitation becomes particularly pronounced in infrared small target detection, where targets are characterized by extremely small spatial scales and weak texture cues, thereby reducing the robustness of such models in complex environments. Motivated by these challenges, recent studies have shifted toward higher-order modeling strategies that explicitly integrate external prior knowledge to better exploit the structural and statistical properties inherent in infrared imagery. For instance, by incorporating non-local autocorrelation characteristics of infrared backgrounds together with sparse representation assumptions, the method in [12] enables accurate separation and detection of small targets. Liu et al. [13] further enhanced detection performance by introducing an effective strategy that accounts for target scale variation and spatial positional information. More recently, Xia et al. [14] proposed an approach that constructs spatiotemporal patch tensors and imposes non-local patch similarity constraints to facilitate infrared data completion and reconstruction. Beyond these approaches, several recent works have explored low-rank tensor modeling for infrared small target detection, aiming to uncover latent non-local correlations within background components. Representative efforts include a Hankel tensor-based subspace representation model combined with non-convex tensor completion techniques, which enables effective recovery of low-rank background structures from infrared images [15].

In this work, infrared small target detection is formulated within a generalized robust principal component analysis (RPCA) framework. A supervised learning strategy is adopted to construct a dictionary that is both discriminative and capable of faithful signal reconstruction. To effectively incorporate label information into the dictionary learning process, an ideal coding regularization term is introduced into the optimization objective. Furthermore, by imposing sparsity constraints on the low-rank formulation, the proposed model is able to learn sparse and structurally meaningful representations of infrared imagery.

The main contributions of this study are summarized as follows:

(i) A novel image representation method based on structured robust principal component analysis is proposed, which enables effective detection of small targets embedded in complex backgrounds.

(ii) A discriminative yet reconstructive dictionary is designed to jointly capture the low-rank background component and the sparse target component of infrared images.

(iii) The proposed algorithm demonstrates strong robustness, maintaining reliable detection performance even in infrared scenes with heavy clutter and extremely weak small target signals.

2. Related works

Many researchers have focused on infrared small-target detection using data structures, where small targets exhibit inherently sparse properties that facilitate their separation. Such detection methods are categorized into single-frame approaches, which analyze individual images, and multi-frame approaches, which exploit temporal information across consecutive frames.

2.1 Single-image infrared small target detection

Recent studies have also explored infrared small target detection from the perspectives of weak supervision and computationally efficient modeling. Ying et al. [16] proposed a label evolution strategy based on single-point annotations, where intermediate predictions generated by a convolutional neural network during training are iteratively exploited to propagate sparse labels, enabling end-to-end pixel-level learning of target masks. To address the issue of insufficient target-background contrast, Zhang et al. [17] drew inspiration from Taylor finite-difference theory and designed an edge-enhancement module coupled with a bidirectional attention aggregation scheme, which effectively emphasizes target boundaries and captures fine-grained shape information. From a semantic modeling viewpoint, Nian et al. [18] integrated gradient-flow-based segmentation, local nonlinear feature representations, and multi-scale feature fusion to significantly strengthen the semantic expressiveness of small targets.

To balance semantic abstraction with detail preservation, Zhang et al. [19] introduced a parallel encoder block (PEB) that combines Transformer and convolutional architectures, and further proposed a stochastic connection attention mechanism to sparsity information flow, thereby reducing model complexity and parameter count. In terms of network light-weighting, Zhang et al. [20] incorporated a wavelet-structured regularization-guided soft channel pruning strategy, achieving improved detection accuracy while substantially decreasing computational cost. In addition, Wu et al. [21] constructed a dedicated dataset for spaceborne infrared dim small vessel detection, providing valuable data support for subsequent research.

With respect to architectural innovations, Zhang et al. [22] redesigned the encoder-decoder structure of segment anything model (SAM) and introduced a Perona-Malik diffusion-based feature enhancement module along with a granularity-aware decoder, leading to superior performance across multiple evaluation metrics. Furthermore, Zhang et al. [23] combined a context-mixing decoding scheme with spatial-frequency attention mechanisms and an eye-shaped enhancement module, resulting in notable improvements in overall infrared small target detection performance.

2.2 Multi-frame infrared small target detection

From the perspective of modeling target scale variations and spatial positional information, Liu et al. [24] proposed an effective strategy to enhance detection performance. Concurrently, a range of emerging network architectures has been applied to infrared small target detection, including Transformer-based models [25], implicit neural representation frameworks [26], and graph neural networks leveraging non-local relational modeling [27]. These methods are designed to build more discriminative representations in high-dimensional feature spaces, thereby improving detection robustness in complex and dynamically changing background environments. Nevertheless, most data-driven methods tend to approximate infrared images using quasi-convex distributions and learn direct mappings from input imagery to segmentation masks. Such formulations often fail to adequately capture the intrinsic structural characteristics of infrared backgrounds. To address this limitation, Xia et al. [28] proposed constructing spatiotemporal patch tensors and imposing non-local patch similarity constraints to facilitate completion and reconstruction of infrared sequence data. Beyond the line of work, recent studies have incorporated low-rank tensor modeling into infrared small target detection to explicitly exploit non-local correlations within background components. For example, a Hankel tensor-based subspace representation model combined with non-convex tensor completion was introduced to recover low-rank structural features from infrared imagery [29].

Building upon these advances, Liu et al. [30] introduced the tensor nuclear norm (TNN) to naturally extend conventional matrix completion frameworks to the tensor domain, where the TNN is defined as a weighted summation of the nuclear norms of unfolded tensor matrices. Mu et al. [31] further reformulated the nuclear norm of square matrices constructed from tensors via re-parameterized modeling. However, existing studies have demonstrated that reconstruction accuracy based on nuclear norm minimization can deteriorate significantly under high sampling ratios along rows or columns, highlighting inherent limitations of such approaches [32].

3. Infrared small target detection technology optimization

3.1 Problem description

Infrared small target detection can be viewed as an optimization problem: perform robust principal component analysis on the input infrared image, i.e.,

$$\boldsymbol{X}=\boldsymbol{L}+\boldsymbol{S} \tag{1}$$

where $\boldsymbol{L}$ is a low-rank matrix and $\boldsymbol{S}$ is a sparse matrix.

Decompose the input $\boldsymbol{X}$ into $\boldsymbol{L}$+$\boldsymbol{S}$ and minimize the rank of $\boldsymbol{L}$:

$$\begin{gathered}\min_{\boldsymbol{L},\boldsymbol{S}} \operatorname{rank}(\boldsymbol{L})+\lambda\|\boldsymbol{S}\|_0\\ \text{s.t. } \boldsymbol{X}=\boldsymbol{L}+\boldsymbol{S}\end{gathered} \tag{2}$$

where λ is a balancing parameter.

The above is an non-deterministic polynomial hard problem. Minimizing the rank of $\boldsymbol{L}$ can be approximated by minimizing the $\boldsymbol{L}$ nuclear norm, which is feasible in mathematical theory. Therefore, the above optimization problem can be transformed into:

$$\begin{gathered}\min_{\boldsymbol{L},\boldsymbol{S}} \|\boldsymbol{L}\|_*+\lambda\|\boldsymbol{S}\|_1\\ \text{s.t. } \boldsymbol{X}=\boldsymbol{L}+\boldsymbol{S}\end{gathered} \tag{3}$$

where $\|\boldsymbol{L}\|_*$ is the nuclear norm of $\boldsymbol{L}$, that is, the sum of singular values; $\|\boldsymbol{S}\|_1$ is the $\boldsymbol{L}_1$ norm of $\boldsymbol{S}$.

The background of infrared images is low-rank compared to infrared small targets; using principal component analysis, the infrared background is decomposed during the alternating optimization of low-rank and sparse components. Consider the infrared small-target detection problem. Here, the dataset is a union of many subjects; samples of one subject tend to be drawn from the same subspace, while samples of different subjects are drawn from different subspaces, as shown in Fig. 1. For the image on the left, a discriminative and reconstructive dictionary is constructed through robust principal component analysis using a supervised learning approach, thereby obtaining the image on the right with the desired characteristics. A more general rank-minimization problem is formulated as

$$\begin{gathered}\min_{\boldsymbol{L},\boldsymbol{S}} \|\boldsymbol{Z}\|_*+\lambda\|\boldsymbol{S}\|_{2,1}\\ \text{s.t. } \boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{S}\end{gathered} \tag{4}$$

where $\boldsymbol{Z}$ is the representation matrix, and $\boldsymbol{D}$ is a dictionary that linearly spans the data space. The quality of $\boldsymbol{D}$ will affect the small-target extraction capability of the representation $\boldsymbol{S}$.

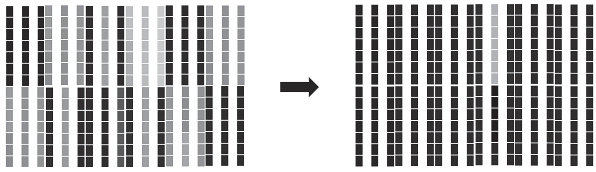


Fig. 1 Construction of the required image

Learn low-rank and sparse representations of images. Low-rankness reveals structural information. Sparsity separates background and small targets. Given a dictionary $\boldsymbol{D}$, the objective function is formulated as

$$\begin{gathered}\min_{\boldsymbol{Z},\boldsymbol{S}} \|\boldsymbol{Z}\|_*+\lambda\|\boldsymbol{S}\|_1+\beta\|\boldsymbol{Z}\|_1\\ \text{s.t. } \boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{S}\end{gathered} \tag{5}$$

where λ, β control the sparsity of the noise matrix $\boldsymbol{S}$ and the representation matrix $\boldsymbol{Z}$, respectively.

Learn a semantically structured dictionary via supervised learning. We add a regularization term $\|\boldsymbol{Z}-\boldsymbol{Q}\|_{\mathrm{F}}^2$ to incorporate structural information into the dictionary learning process, where $\boldsymbol{Q}$ is the ideal representation of $\boldsymbol{Z}$. The dictionary learning objective is defined as

$$\begin{gathered}\min_{\boldsymbol{Z},\boldsymbol{S},\boldsymbol{D}} \|\boldsymbol{Z}\|_*+\lambda\|\boldsymbol{S}\|_1+\beta\|\boldsymbol{Z}\|_1+\alpha\|\boldsymbol{Z}-\boldsymbol{Q}\|_{\mathrm{F}}^2\\ \text{s.t. } \boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{S}\end{gathered} \tag{6}$$

where α is the weight controlling the regularization term.

3.2 Optimization procedure

To solve the proposed optimization problem, an auxiliary variable $\boldsymbol{W}$ is first introduced to make the objective separable. The problem can be rewritten as

$$\begin{gathered}\min_{\boldsymbol{Z},\boldsymbol{S},\boldsymbol{D}} \|\boldsymbol{Z}\|_*+\lambda\|\boldsymbol{S}\|_1+\beta\|\boldsymbol{W}\|_1+\alpha\|\boldsymbol{W}-\boldsymbol{Q}\|_{\mathrm{F}}^2\\ \text{s.t.} \begin{cases}\boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{S}\\ \boldsymbol{W}=\boldsymbol{Z}\end{cases}\end{gathered} \tag{7}$$

where $\boldsymbol{W}$ is a auxiliary variable.

Its augmented Lagrangian function is

$$\begin{gathered}L(\boldsymbol{Z},\boldsymbol{W},\boldsymbol{S},\boldsymbol{D},\boldsymbol{Y}_1,\boldsymbol{Y}_2,\mu)=\\ \|\boldsymbol{Z}\|_*+\lambda\|\boldsymbol{S}\|_1+\beta\|\boldsymbol{W}\|_1+\alpha\|\boldsymbol{W}-\boldsymbol{Q}\|_{\mathrm{F}}^2+\\ h(\boldsymbol{Z},\boldsymbol{W},\boldsymbol{S},\boldsymbol{D},\boldsymbol{Y}_1,\boldsymbol{Y}_2,\mu)-\frac{1}{2\mu}\left(\|\boldsymbol{Y}_1\|_{\mathrm{F}}^2+\|\boldsymbol{Y}_2\|_{\mathrm{F}}^2\right)\end{gathered} \tag{8}$$

where

$$\begin{gathered}h(\boldsymbol{Z},\boldsymbol{W},\boldsymbol{S},\boldsymbol{D},\boldsymbol{Y}_1,\boldsymbol{Y}_2,\mu)=\\ \frac{\mu}{2}\left(\left\|\boldsymbol{X}-\boldsymbol{D}\boldsymbol{Z}-\boldsymbol{S}+\frac{\boldsymbol{Y}_1}{\mu}\right\|_{\mathrm{F}}^2+\left\|\boldsymbol{Z}-\boldsymbol{W}+\frac{\boldsymbol{Y}_2}{\mu}\right\|_{\mathrm{F}}^2\right),\end{gathered} \tag{9}$$

$\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are two multiplier terms, and μ and η are two equilibrium coefficients.

Minimize this function by alternately updating the variables $\boldsymbol{Z}$, $\boldsymbol{W}$, $\boldsymbol{S}$. The specific steps are as follows:

Update $\boldsymbol{Z}^{j+1}$:

$$\boldsymbol{Z}^{j+1}=\arg\min\frac{1}{\eta\mu}\|\boldsymbol{Z}\|_*+\frac{1}{2}\left\|\boldsymbol{Z}-\boldsymbol{Z}^j+\frac{\left[-\boldsymbol{D}^{\mathrm{T}}\left(\boldsymbol{X}-\boldsymbol{D}\boldsymbol{Z}^j-\boldsymbol{S}^j+\boldsymbol{Y}_1^j/\mu\right)+\left(\boldsymbol{Z}-\boldsymbol{W}^j+\boldsymbol{Y}_2^j/\mu\right)\right]}{\eta}\right\|_{\mathrm{F}}^2. \tag{10}$$

Update $\boldsymbol{W}^{j+1}$:

$$\boldsymbol{W}^{j+1}=\arg\min_{w}\frac{\beta}{2\alpha+\mu}\|\boldsymbol{W}\|_1+\frac{1}{2}\left\|\boldsymbol{W}-\left(\frac{2\alpha}{2\alpha+\mu}\boldsymbol{Q}^{+}\frac{1}{2\alpha+\mu}\boldsymbol{Y}_2^j+\frac{\mu}{2\alpha+\mu}\boldsymbol{Z}^{j+1}\right)\right\|_{\mathrm{F}}^2. \tag{11}$$

Update $\boldsymbol{S}^{j+1}$:

$$\boldsymbol{S}^{\mathrm{j}+1}=\arg\min_{s}\frac{\lambda}{\mu}\|\boldsymbol{S}\|_1+\frac{1}{2}\left\|\boldsymbol{S}-\left(\frac{1}{\mu}\boldsymbol{Y}_1^{\mathrm{j}}+\boldsymbol{X}-\boldsymbol{D}\boldsymbol{Z}^{j+1}\right)\right\|_{\mathrm{F}}^2. \tag{12}$$

By the above alternating direction method of multipliers, the small targets in infrared images can be ultimately solved by iteration.

4. Experimental validation

To objectively demonstrate the superiority of our algorithm, we compare it with several advanced methods on a public infrared small-target detection dataset. The compared methods are the top-performing algorithms from the past two years. These methods involve different network architectures, such as convolutional neural networks (CNN) and Transformers. The compared algorithms include graph-based context learning network (GCLNet) [33] (2025), spatial-channel cross transformer network (SCTransNet) [34] (2024), dense nested attention network (DNA-Net) [35] (2023), and attention-guided pyramid context network (AGPCNet) [36] (2023). All our training, testing, and validation experiments are performed on a host equipped with an i9 CPU and 64 GB RAM.

4.1 Quantitative comparison experiments

The quantitative comparison experiments are conducted on the commonly used public dataset Infrared Small Target Detection-1k (IRSTD-1k). Comparison metrics include detection rate (p_{d}), false alarm rate (F_{a}), intersection over union (IOU), normalized IOU (nIOU), and the F_1 measure. The formulas for these five metrics are as follows.

Detection rate:

$$P_{\mathrm{d}}=\frac{N_{\mathrm{pred}}}{N_{\mathrm{all}}} \tag{13}$$

where N_{pred} is the number of correctly predicted targets; N_{all} is the total number of targets.

False alarm rate:

$$F_{\mathrm{a}}=\frac{N_{\mathrm{false}}}{N_{\mathrm{all}}} \tag{14}$$

where N_{false} is the number of false positive targets.

nIOU:

$$\mathrm{nIOU}=\frac{A_i}{A_u}=\frac{\sum_{i=1}^{N}\mathrm{TP}[i]}{\sum_{i=1}^{N}(\mathrm{T}[i]+\mathrm{P}[i]-\mathrm{TP}[i])} \tag{15}$$

where TP (true positive) represents the number of pixels (or area) where the predicted region overlaps with the true target region, T (ground truth target) denotes the area of the true labeled target region (the total number of true target pixels), P (prediction) signifies the area of the predicted target region (the total number of target pixels predicted by the algorithm), A_i (intersection area) indicates the intersection area between the predicted region and the true region, A_u (union area) represents the union area of the predicted region and the true region.

F_1 metric:

$$F_1=\frac{2\mathrm{Prec}\cdot\mathrm{Rec}}{\mathrm{Prec}+\mathrm{Rec}} \tag{16}$$

where Prec and Rec are precision and recall, respectively.

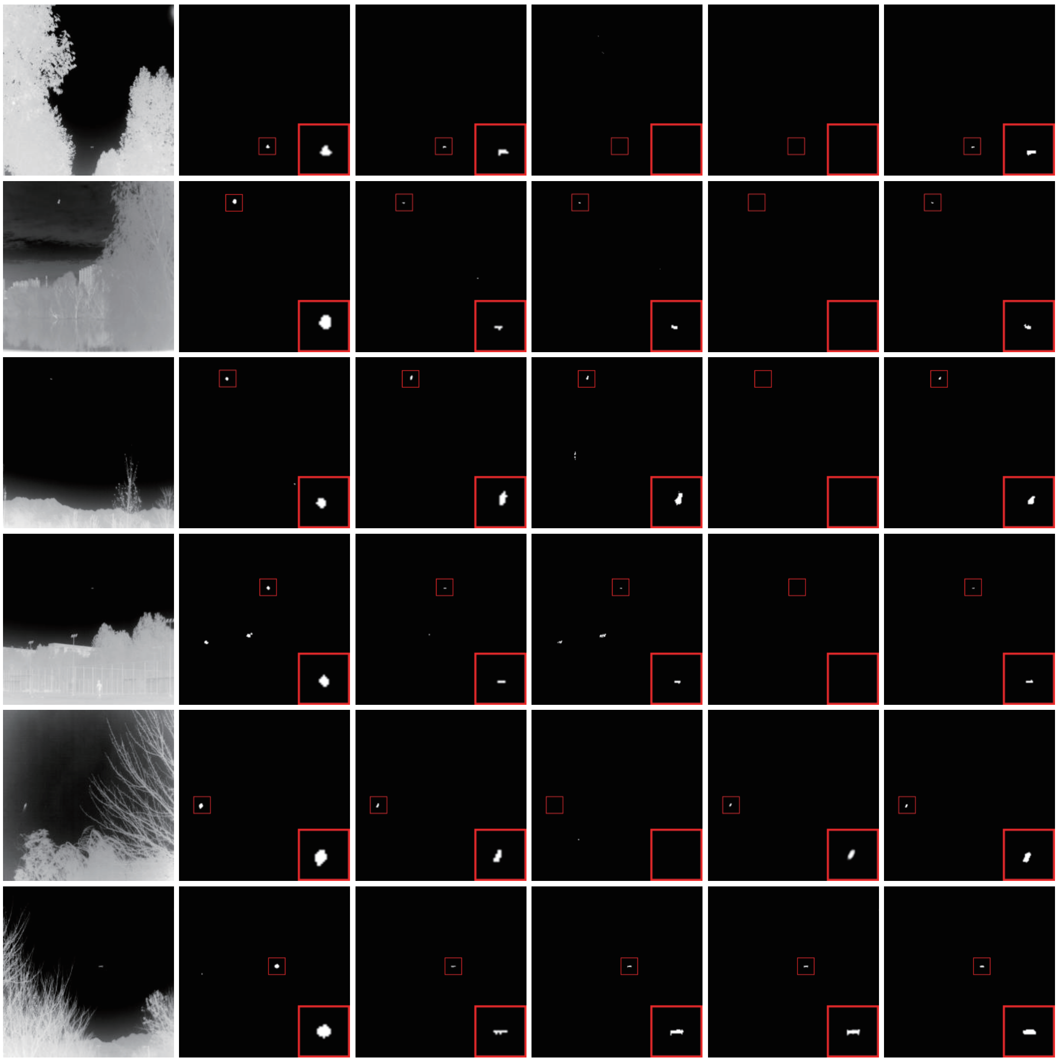
Table 1 shows quantitative comparison results between the proposed algorithm and four state-of-the-art infrared small-target detection algorithms, including GCLNet, SCTransNet, DNA-Net, and AGPCNet. Among them, the underlined data represents suboptimal results, and the bolded data represents the optimal results. The results indicate that the proposed robust principal component analysis-based infrared small-target detection algorithm achieves excellent quantitative results, with the highest detection rate among all compared algorithms. SCTransNet, based on the Transformer framework, also performes well, achieving the lowest false alarm rate. AGPCNet's results are unsatisfactory; its designed network structure has significant flaws, putting it at a disadvantage across all metrics. The attention mechanism built in DNA-Net effectively aggregates target features, thereby successfully detecting small targets.

Table 1 Quantitative comparison experiment results conducted on the public dataset IRSTD-1k

Algorithm	IRSTD-1k				
	IOU	nIOU	F_1	P_{d}	F_{a}
GCLNet [33]	<u>81.45</u>	**82.31**	<u>91.32</u>	<u>97.26</u>	<u>17.26</u>
SCTransNet [34]	80.26	80.78	87.16	90.25	**13.25**
DNA-Net [35]	76.33	77.51	83.10	91.64	31.00
AGPCNet [36]	63.25	65.07	76.81	95.20	36.25
Proposed method	**81.56**	<u>82.30</u>	**92.56**	**98.21**	17.77

4.2 Qualitative comparison experiments

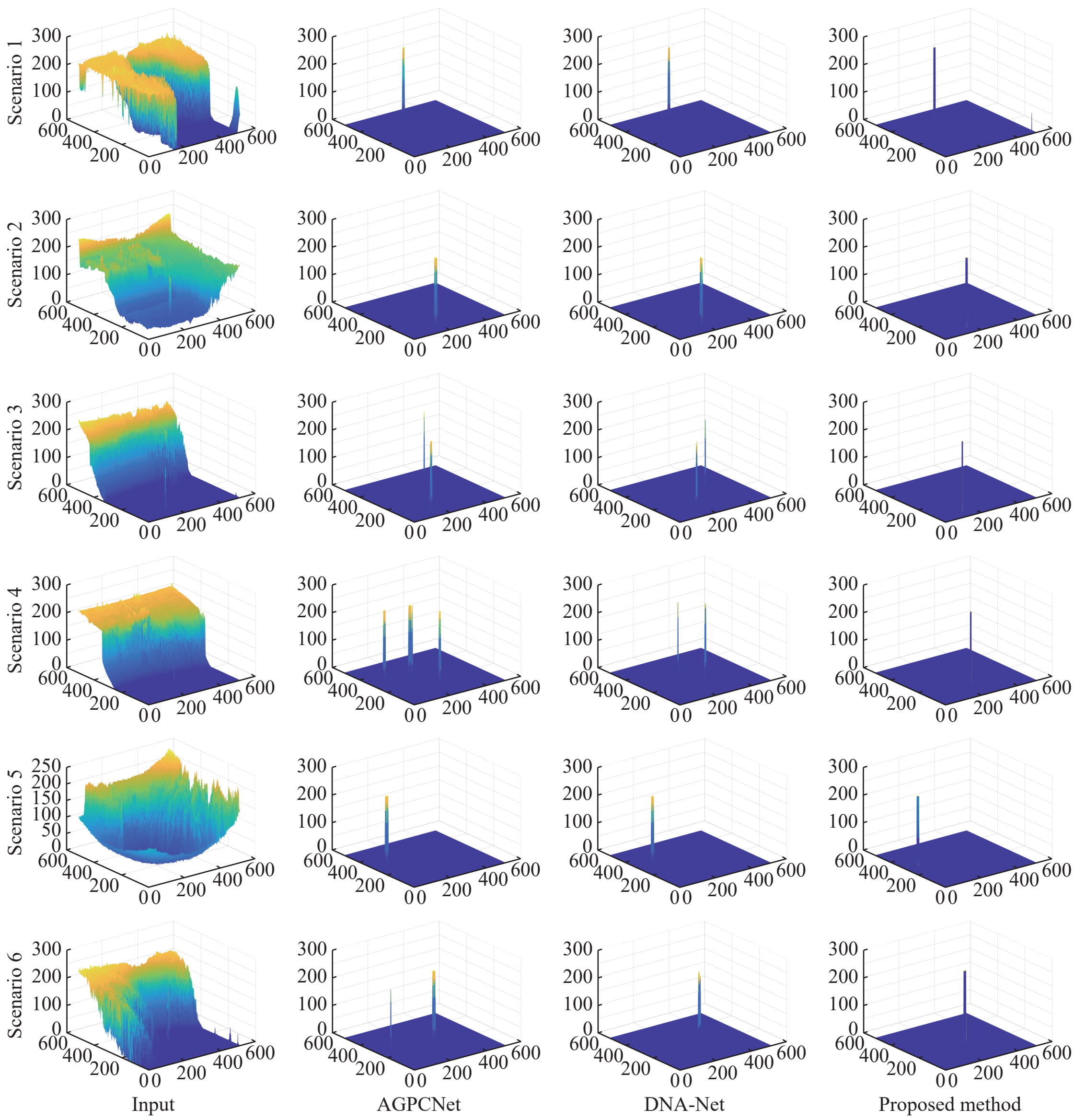
In addition to quantitative comparison experiments, qualitative comparisons are also conducted. Fig. 2 shows the results of qualitative comparison experiments. From these results, it can be seen that across diverse test scenarios, all compared algorithms performed fair detection. Even in the face of complex backgrounds, it can still effectively eliminate clutter. The six displayed scenes are varied, including aerial targets and complex backgrounds. Scenes involve branches, buildings, artificial streetlights, pedestrians, and so on. The scenes are very complex and representative, and are the kinds of situations infrared cameras often face in crowded or wild areas.



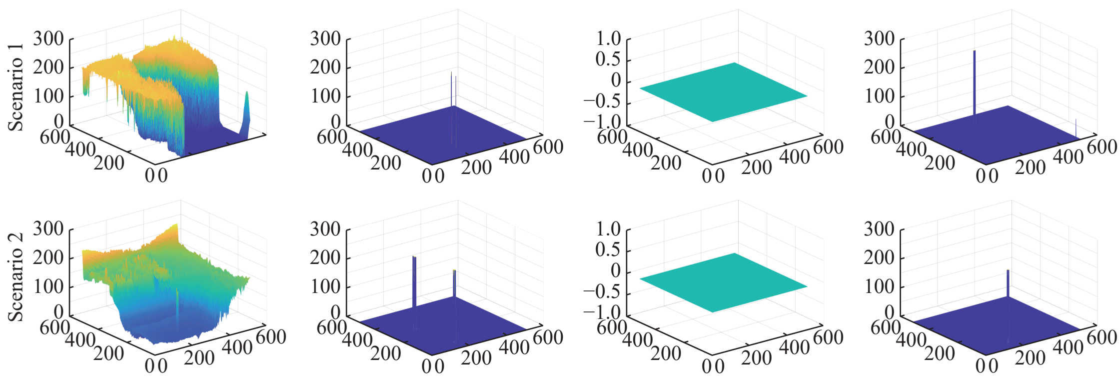
(a) Input (b) AGPCNet (c) DNA-Net (d) GCLNet (e) SCTransNet (f) Proposed method

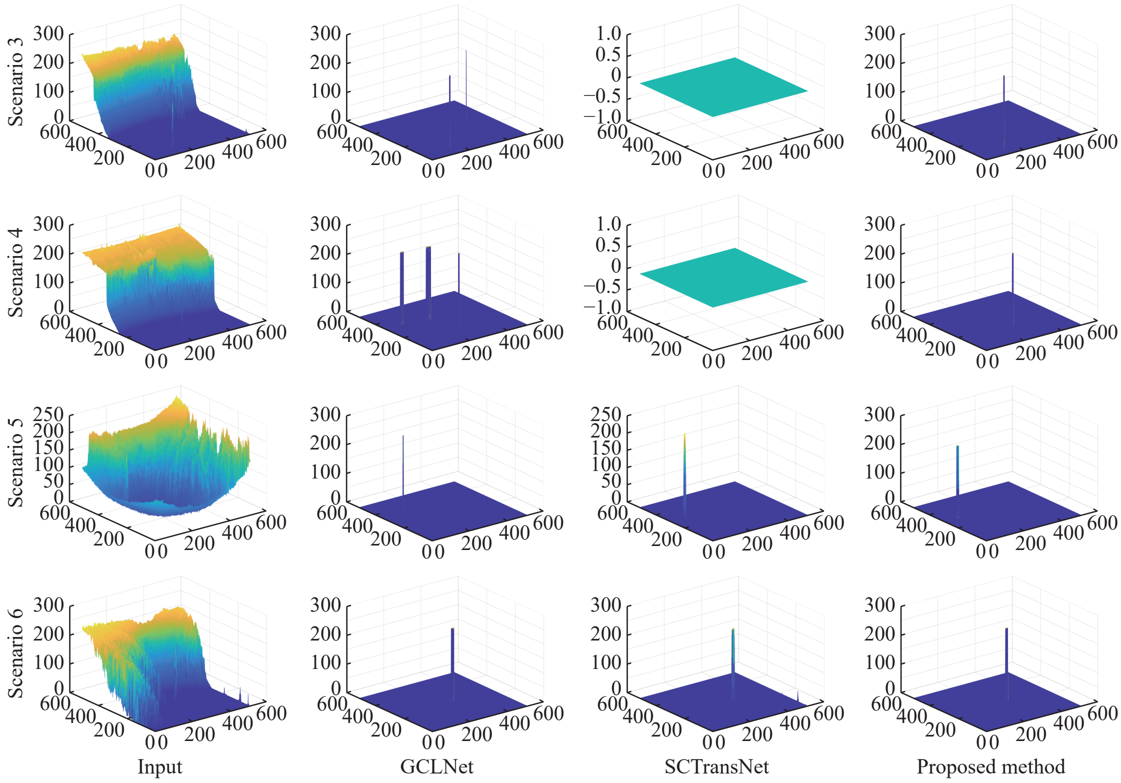
Fig. 2 Experimental results in typical scenarios

In the comparison results, although the AGPCNet algorithm detects all targets, its false alarm rate is high. Many results contain background information because its network's threshold for distinguishing background and targets is low and its capability is weak. SCTransNet's detection results are unsatisfactory; in many common scenes it almost fails to detect targets, and even though it has the lowest false alarm rate, that does not help detection much. DNA-Net's results are relatively good, but there are still many visible clutters, as can be seen in the results of Fig. 3. For handling clutter, SCTransNet is not as strong as GCLNet; GCLNet achieves impressive experimental results and a decent detection rate. Even so, it is still below the algorithm we propose. Our proposed algorithm is unique in handling clutter and noise: robust principal component analysis can accurately extract small-target features and summarize background information. The results show that the algorithm proposed in this paper detects all small targets, and even when faced with complex backgrounds, the detection results are still the best.



(a) Comparison of AGPNet algorithm and DNA-Net algorithm





(b) Comparison of GCLNet algorithm and SCTransNet algorithm

Fig. 3 Three-dimensional intensity maps of six scenarios and the detection results of all algorithms

5. Conclusions

This paper proposes an effective robust principal component analysis model to address the small-target detection problem in complex scenes. To tackle the difficulty of extracting small-target features in complex scenes and their susceptibility to interference, we propose an algorithmic framework based on robust principal component analysis specifically designed for infrared small-target detection in complex backgrounds. First, we formulate infrared small-target detection as a general robust principal component analysis process and construct a discriminative learned dictionary. Second, by introducing regularization terms, we reconstruct infrared images without losing structural information to obtain discriminative principal component representations relative to the constructed dictionary. Finally, we reconstruct the structural features of infrared small targets via sparse coding to extract the infrared small targets. In future work, we will continue to optimize the robust principal component analysis model and design more appropriate regularization terms and dictionaries to improve detection results.

References

[1] HADHOUD M M, THOMAS D W. The two-dimensional adaptive LMS (TDLMS) algorithm. IEEE Trans. on Circuits and Systems, 1988, 35(5): 485–494.

[2] SONI T, ZEIDLER J R, KU W H. Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data. IEEE Trans. on Image Processing, 1993, 2(3): 327–340.

[3] HAN J H, YONG M, BO Z, et al. A robust infrared small target detection algorithm based on human visual system. IEEE Geoscience and Remote Sensing Letters, 2014, 11(12): 2168–2172.

[4] WEI Y T, YOU X G, LI H. Multiscale patch-based contrast measure for small infrared target detection. Pattern Recognition, 2016, 58: 216–226.

[5] LI Y J, WANG L P, CHEN S C. Smile: spatial-spectral mamba interactive learning for infrared small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2025, 63: 5005214.

[6] TONG Y F, LIU J, FU Z L, et al. Guided attention and joint loss for infrared dim small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2024, 62: 4110914.

[7] AIBIBU T, LAN J H, ZENG Y L, et al. An efficient rep-style Gaussian-Wasserstein network: improved UAV infrared small object detection for urban road surveillance and safety. Remote Sensing, 2023, 16(1): 25.

[8] GIRSHICK R. Fast R-CNN. Proc. of the IEEE International Conference on Computer Vision, 2015: 1440−1448.

[9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection. Proc. of the IEEE Conference on Computer Vision and Pattern Recogni-

tion, 2016: 779−788.

[10] CHEN Z H, WU G Y, GAO H M, et al. Local aggregation and global attention network for hyperspectral image classification with spectral-induced aligned superpixel segmentation. Expert Systems with Applications, 2023, 232: 120828.

[11] YANG B, CHENG X W, CHEN W, et al. A graph-based hyperspectral change detection framework using difference augmentation and progressive reconstruction with limited labels. IEEE Trans. on Geoscience and Remote Sensing, 2024, 62: 5518914.

[12] KONG X, YANG C P, CAO S Y, et al. Infrared small target detection via nonconvex tensor fibered rank approximation. IEEE Trans. on Geoscience and Remote Sensing, 2022, 60: 5000321.

[13] JI S, ZHANG H F, ZHANG J G, et al. A three-stage model for infrared small target detection with spatial and semantic feature fusion. Expert Systems with Applications, 2025, 295: 128776.

[14] LI Y J, WANG L P, CHEN S C. From optimization to network: a low-rank and sparse-aware deep unfolding framework for infrared small target detection. Advanced Engineering Informatics, 2026, 69: 103991.

[15] LI Y J, WANG L P, CHEN S C. Frequency-spatial interaction reinforcement paradigm for infrared small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2025, 63: 5008915.

[16] YING X Y, LIU L, WANG Y Q, et al. Mapping degeneration meets label evolution: learning infrared small target detection with single point supervision. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 15528−15538.

[17] ZHANG M J, ZHANG R, YANG Y X, et al. ISNet: shape matters for infrared small target detection. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 867−876.

[18] NIAN B K, JIANG B, SHI H J, et al. Local contrast attention guide network for detecting infrared small targets. IEEE Trans. on Geoscience and Remote Sensing, 2023, 61: 5607513.

[19] ZHANG M J, BAI H C, ZHANG J, et al. RKformer: Runge-Kutta transformer with random-connection attention for infrared small target detection. Proc. of the 30th ACM International Conference on Multimedia, 2022: 1730−1738.

[20] ZHANG M J, YANG H D, GUO J, et al. IRPruneDet: efficient infrared small target detection via wavelet structure-regularized soft channel pruning. Proc. of the AAAI Conference on Artificial Intelligence, 2024, 38(7): 7224–7232.

[21] WU T H, LI B Y, LUO Y H, et al. MTU-Net: multilevel transunet for space-based infrared tiny ship detection. IEEE Trans. on Geoscience and Remote Sensing, 2023, 61: 5601015.

[22] ZHANG M J, WANG Y C, GUO J, et al. IRSAM: advancing segment anything model for infrared small target detection. Proc. of the 18th European Conference on Computer Vision, 2024: 233−249.

[23] ZHANG M J, ZHANG R, ZHANG J, et al. Dim2Clear network for infrared small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2023, 61: 5626015.

[24] LIU Q K, LIU R, ZHENG B L, et al. Infrared small target detection with scale and location sensitivity. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 17490−17499.

[25] CHOI L, CHUNG W Y, PARK C G. CSI-Net: CNN Swin transformer integrated network for infrared small target detection. International Journal of Control, Automation and Systems, 2024, 22(9): 2899–2908.

[26] LIU P, LUO Y S, WANG W Z, et al. Motion-enhanced nonlocal similarity implicit neural representation for infrared dim and small target detection. https://arxiv.org/abs/2504.15665.

[27] JIA G M, CHENG Y, CHEN T. IRGraphSeg: infrared small target detection based on hierarchical GNN. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 6005505.

[28] XIA C Q, CHEN S H, HUANG R H, et al. Separable spatial-temporal patch-tensor pair completion for infrared small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2024, 62: 5001620.

[29] MA F, QU Q, YANG F X, et al. Hankel tensor subspace representation for remotely sensed image fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 8630–8642.

[30] LIU J, MUSIALSKI P, WONKA P, et al. Tensor completion for estimating missing values in visual data. TIEEE Trans. on Pattern Analysis and Machine Intelligence, 2013, 35(1): 208–220.

[31] MU C, HUANG B, WRIGHT J, et al. Square deal: lower bounds and improved relaxations for tensor recovery. Proc. of the 31st International Conference on Machine Learning, 2014, 32(2): 73−81.

[32] SALAKHUTDINOV R, SREBRO N. Collaborative filtering in a non-uniform world: learning with the weighted trace norm. Proc. of the 24th International Conference on Neural Information Processing Systems, 2010: 2056−2064.

[33] SHEN Y W, LI Q W, XU C, et al. Graph-based context learning network for infrared small target detection. Neurocomputing, 2025, 616: 128949.

[34] YUAN S, QIN H L, YAN X, et al. SCTransNet: spatial-channel cross transformer network for infrared small target detection. IEEE Trans. on Geoscience and Remote Sensing, 2024, 62: 5002615.

[35] LI B Y, XIAO C, WANG L G, et al. Dense nested attention network for infrared small target detection. IEEE Trans. on Image Processing, 2022, 32: 1745–1758.

[36] ZHANG T Y, LI L, CAO S Y, et al. Attention-guided pyramid context networks for detecting infrared small target under complex background. IEEE Trans. on Aerospace and Electronic Systems, 2023, 59(4): 4250–4261.