remote sensing

MDPI

Article

ILN-SSR: Improved Logarithmic Norm and Sparse Structure Refinement for Infrared Small Target Detection

Liqi Liu [1,2,3], Rongguo Zhang [1,2,3], Jian Mei [1,2,3], Xinyue Ni [1,2,*], Liyuan Li [4], Xiaofeng Su [1,2] and Fansheng Chen [1,2,4]

[1] Shanghai Institute of Technical Physics of the Chinese Academy of Sciences, Shanghai 200083, China; liuliqi@mail.sitp.ac.cn (L.L.); zhangrongguo22@mails.ucas.ac.cn (R.Z.); meijian@mail.sitp.ac.cn (J.M.); fishsu@mail.sitp.ac.cn(X.S.); cfs@mail.sitp.ac.cn (F.C.)
[2] Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences, Shanghai 200083, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[4] Shanghai Frontier Base of Intelligent Optoelectronics and Perception, Institute of Optoelectronics, Fudan University, Shanghai 200433, China; liliyuan@mail.sitp.ac.cn
* Correspondence: nixinyue@mail.sitp.ac.cn

Abstract: The effective discrimination of targets from backgrounds in environments characterized by a low signal-to-clutter ratio (SCR) is paramount for the advancement of infrared small target detection (IRSTD). In this work, we propose a novel detection framework predicated on low-rank sparse decomposition (LRSD), incorporating an improved logarithmic norm and a mechanism for sparse structure refinement, herein referred to as the improved logarithmic norm and sparse structure refinement (ILN-SSR). The ILN-SSR framework more precisely characterizes the sparse properties of both the background and the target, enabling a more effective distinction between the target and its background. Initially, our approach entails the utilization of an improved logarithmic norm to precisely estimate the low-rank attributes of the infrared image background. This is followed by the employment of a linear sparse regularization term alongside a target-traits-based sparse regularization term aimed at meticulously identifying targets within sparse regions and refining the sparse structure. Subsequently, we combine these components into the ILN-SSR framework, which formulates IRSTD as an optimization problem. The resolution of this framework is achieved through the implementation of the alternating direction method of multipliers (ADMM). The efficacy of the proposed framework is corroborated through the analysis of six image sequences. Comprehensive experimental assessments affirmed the framework's substantial robustness in navigating various complex backgrounds.

Keywords: alternating direction method of multipliers (ADMM); infrared small target detection (IRSTD); improved logarithmic norm and sparse structure refinement (ILN-SSR); low-rank sparse decomposition (LRSD); signal-to-clutter ratio (SCR)

check for updates

Citation: Liu, L.; Zhang, R.; Mei, J.; Ni, X.; Li, L.; Su, X.; Chen, F. ILN-SSR: Improved Logarithmic Norm and Sparse Structure Refinement for Infrared Small Target Detection. *Remote Sens.* **2024**, *16*, 4018. https://doi.org/10.3390/rs16214018

Academic Editors: Hao Li, Mingyang Zhang and Gonzalo Pajares Martinsanz

Received: 22 August 2024
Revised: 19 October 2024
Accepted: 24 October 2024
Published: 29 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In contemporary applied research, infrared imaging systems have emerged as a prominent subject of investigation, assuming a paramount role within domains associated with national defense security [1–4]. Within this realm, infrared small target detection (IRSTD) occupies a pivotal position. In accordance with the guidelines established by the Society of Photo-Optical Instrumentation Engineers (SPIE), a small infrared target is characterized as one that either does not surpass a dimension of 9×9 pixels or occupies a fractional area of less than 0.15% of the total image [5,6]. Furthermore, in practical infrared detection scenarios, small targets typically exhibit weak brightness and lack sufficient textural information. These specific types of targets exhibit a tendency to become obscured within intricate and fluctuating environmental backgrounds, frequently composed of elements such as buildings, roads, atmospheric formations (e.g., clouds), ground clutter, and additional constituents. In scenarios characterized by a low signal-to-clutter ratio (SCR), prevailing

detection methodologies frequently encounter challenges in effectively discriminating between targets and the background. This deficiency often results in misdetections or false alarms, thereby compromising the reliability and accuracy of the detection process. Consequently, achieving robust separation of infrared targets from the background in environments characterized by low signal-to-clutter ratios persists as a formidable challenge.

The existing IRSTD methods can be summarized into the following categories: methods based on background filtering, methods based on the human visual system (HVS), methods based on deep learning, and methods based on low-rank sparse decompositions (LRSDs). Detection methods based on background filtering distinguish targets from backgrounds by analyzing statistical correlations and utilizing spatial [7–10] and frequency [11–13] domain techniques. Although effective, these methods often face stability issues in complex scenarios, especially under low SCR conditions. Similarly, human visual system-inspired methods show excellent adaptability across various scenes [14–21] but struggle in complex environments or under intense illumination, often misclassifying structurally similar targets as background, leading to false alarms. Deep learning approaches, which detect targets based on learned textures and dispersion patterns [22–30], also have limitations, particularly when encountering targets without distinct textures or structures.

In the study by Gao et al. [31], the LRSD approach was integrated into the IRSTD framework, where the image background is modeled as a low-rank matrix and the target as a sparse matrix. This conceptualization, known as the Infrared Patch-Image (IPI) model, reformulates the problem of infrared small target detection into a low-rank sparse decomposition (LRSD) problem, which is then addressed using matrix optimization techniques. The IPI model estimates the background using the nuclear norm and estimates the sparse target using the L1 norm, allowing it to adapt effectively to a wide variety of scenes. However, due to the difficulty in accurately characterizing the low-rank properties of the background with the nuclear norm, and because the sparse components cannot be considered equivalent to the target, this model often encounters challenges in practical detection, such as incomplete separation of the target and background or clutter being mistaken for the target, resulting in false alarms and misdetections.

Building on the LRSD-based method described above, this paper addresses the issues of robustly separating targets from the background to enhance detection performance. We propose methods for background estimation and sparse structure refinement. By providing a more accurate description of the background and sparse targets in infrared images, the proposed method can robustly separate targets from the background in complex backgrounds, strong illumination, low-contrast conditions, and scenarios where the target lacks texture. As a result, it effectively reduces false alarms and improves the detection rate.

1.1. Related Works

To enhance the robustness of IRSTD in intricate environments and mitigate diverse forms of interference, scholars have engaged in comprehensive research aimed at identifying weak infrared small targets over the past several decades.

1.1.1. Background Filtering-Based Detection Methods

Detection based on background filtering processes infrared imagery using spatial and frequency domain filters designed to exploit the background's positional and grayscale characteristics. These filters primarily suppress background signals to isolate infrared targets, leveraging the statistical correlation of the background's spatial continuity and symmetry. Spatial domain techniques include TDLMS [7], Max-mean, median filters [8], and morphological filters [9], and random walker [10], while frequency domain methods utilize wavelet transformations [11,12] and Fourier transforms [13].

However, background filtering methods often encounter stability issues in complex scenarios, particularly under low signal-to-clutter ratios, resulting in increased false detections within noisy backgrounds.

1.1.2. HVS-Based Detection Methods

Owing to the robust capabilities of the human visual system in target detection, Kim et al. [14] introduced a novel approach using the Laplacian of Gaussian operator. Building on this, Chen et al. [15] proposed the local contrast method. To address the detection of small targets of varying sizes in complex scenes, enhanced local contrast methods were developed, including NLCM [16], RLCM [17], MPCM [18], TLLCM [19], HWLCM [20], and DNGM [21].

Unlike background filtering methods, these approaches exhibit excellent scene adaptability. However, in complex environments or under intense illumination, they often misclassify structurally similar objects, leading to false alarms.

1.1.3. Deep Learning-Based Detection Methods

The advent of end-to-end learning has enabled deep neural networks to autonomously extract features from complex datasets. Wang et al. [22] used Generative Adversarial Networks to balance missed detections and false alarms, while Dai et al. [23,24] introduced infrared small target datasets and integrated layer expansion, cascading, and segmentation into IRSTD via ALCNet and Oscar. Wu et al. [25] built a multiscale robust IRSTD network by embedding nested Unet structures, originally from semantic segmentation. Li et al. [26,27] developed a denoising and detection network for infrared remote sensing, addressing single-frame limitations by incorporating tracking. Tong et al. [28] leveraged multiscale labels to capture target details, while Zhang et al. [29] utilized Gaussian curvature to reduce speckle noise in detection. Kou's survey [30] evaluated various small target segmentation networks, offering a comprehensive review of deep learning in IRSTD.

However, the success of data-driven methods depends heavily on large training datasets, which are difficult to obtain for infrared small targets. Additionally, current networks lack robustness in varying conditions, highlighting an area where traditional methods still hold advantages in detection accuracy.

1.1.4. LRSD-Based Detection Methods

Recognizing the limitations of the nuclear norm in the IPI model [31] with regard to accurately estimating the image rank, Zhang et al. [32] adopted a strategy wherein some of the larger singular values, determined through the nuclear norm, were omitted, and the partial nuclear norm was proposed as a substitute for rank estimation. In a complementary vein, Wang et al. [33] introduced the concept of the tensor truncated nuclear norm to confine the predominant singular values. Guan et al. [34] harnessed the Laplacian function for smoothing purposes and judiciously assigned adaptive weights to each singular value, while Zhu et al. [35] employed the Log operator to estimate the non-local background within the constructed image.

Furthermore, certain sources of noise within the image continue to pose formidable challenges in the context of detection. To address this concern, Liu et al. [36] incorporated a regularization term involving asymmetric total variation to ameliorate background irregularities in both temporal and spatial domains. Concurrently, Zhang et al. [37] leveraged the non-local autocorrelation characteristics of the image background, proposing a tensor-based non-local total variation (NLTV) method to address sparse clutter, such as edges, manifesting within the background. Drawing inspiration from prior research efforts [38,39], Luo et al. [40] enhanced the l1,1,2-norm and introduced a tubewise sparse regularization term. In a related vein, Kong et al. [41] achieved commendable results by employing hypertotal variation within the context of the infrared patch tensor model.

Despite these commendable advancements, current methodologies for detection still contend with the inherent complexities associated with the thorough decomposition of the background, target, and clutter to a certain extent.

1.2. Motivations and Contributions

In the context of LRSD-based infrared small target detection, the non-convex estimation of image background rank presents a significant challenge. Since directly optimizing the matrix rank is NP-hard, rank optimization is often approached through its convex envelope, the nuclear norm, which is defined as the sum of the matrix's singular values. When the matrix rank is extended to three-dimensional tensors, the corresponding rank, known as the tubal rank, is also NP-hard. The singular values of infrared image backgrounds typically follow a long-tail distribution, where a small number of singular values are relatively large, while the majority of them are small. An efficacious rank approximation function, therefore, must fulfill two pivotal criteria:

(1) Balance the magnitudes of larger singular values through the application of reduced weighting;
(2) Impose greater penalties on nearly zero singular values.

The logarithmic function effectively embodies these desired characteristics, with a steep slope near zero, resulting in a strong penalty on singular values close to zero while balancing the impact on larger singular values. The work in [42] proposed an improved logarithmic norm, which inherits and enhances these properties. It has already been applied in the fields of denoising [43] and instance segmentation [44]. Inspired by these applications, this paper introduces the improved logarithmic norm (ILN) for use in its specific context. Meanwhile, in order to better describe the low-rank properties of the background in optimization, we choose to use a tensor composed of image blocks rather than the original image as the optimization input, as shown in Figure 1a,d.

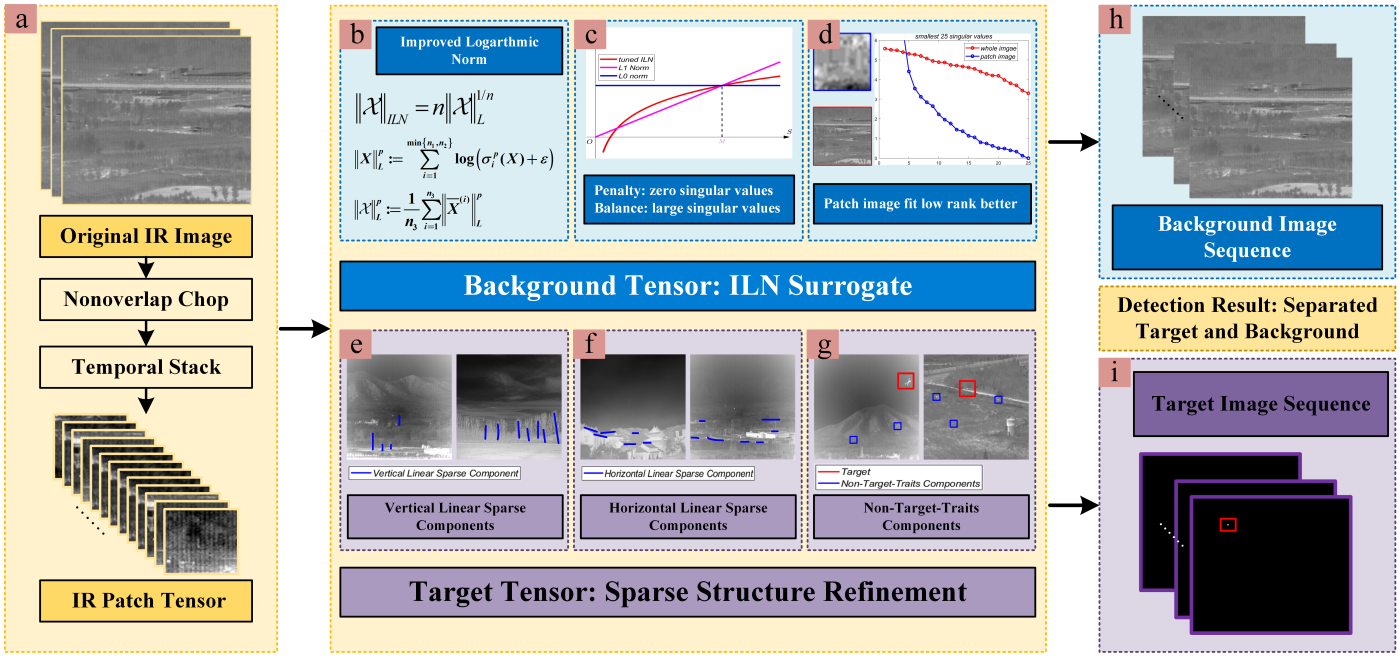


Figure 1. The conceptual diagram of the proposed framework. (**a**) illustrates the process of transforming the original image into a patch tensor. (**b**–**g**) succinctly summarize the innovations presented in this paper, encompassing the improved logarithmic norm and sparse structure refinement. (**h**,**i**) signifies the output of the detection results.

In complex background scenarios, an abundance of clutter elements presents a tough challenge, complicating the task of distinguishing them from the target, thus making it both challenging and time-consuming. To mitigate this challenge, our study proposes the introduction of a local prior weight map to emphasize crucial optimization areas.

Additionally, although the target matrix or tensor is conceptualized as sparse, the practical reality often involves sparse components that encompass not only the target but also robust edges, corners, and sporadic noise artifacts originating from the imaging system. It is evident that classifying a sparse tensor inclusive of these diverse components as the sole representation of the target is a mischaracterization. Previous research, as documented in [39], has successfully employed the $l_{1,1,2}$-norm to eliminate linear noise from video sequences.

Expanding upon this foundational framework, we have formulated linear sparse regularization terms and introduced target-traits-based sparse regularization terms. These innovative components effectively perform sparse structure refinement, as shown in Figure 1e–g. It is important to note that single-frame images lack temporal information, whereas image sequences often contain interconnected data that can greatly enhance detection capabilities.

To address these challenges, we propose a novel IRSTD framework, termed improved logarithmic norm and sparse structure refinement (ILN-SSR). This framework is a target detection methodology built upon the LRSD optimization approach and integrates an improved logarithmic norm with sparse structure refinement techniques. This ILN-SSR is specifically designed to effectively estimate image backgrounds, suppress and eliminate clutter within the sparse data, and robustly detect infrared small targets. The conceptual diagram of the proposed method is illustrated in Figure 1. The key contributions of this paper are encapsulated as follows:

(1) An improved logarithmic norm has been introduced as a method for estimating the low-rank properties of the image background in IRSTD. This approach results in the enhanced preservation of target details, concurrently mitigating the impact of noise.
(2) Linear sparse regularization terms and target-traits-based sparse regularization terms have been introduced, which effectively contribute to sparse structure refinement (SSR), simultaneously filtering out extraneous clutter.
(3) We formulate an IRSTD framework, ILN-SSR, that integrates joint spatial-temporal information for infrared small target detection. We transform the problem of infrared small target detection into an LRSD problem and elucidate the solution approach using an optimization method based on alternating direction method of multipliers (ADMM).

The remainder of this paper is structured as follows. In Section 2, we provide a concise overview of the mathematical symbols employed in tensor algebra. In Section 3, we present the ILN-SSR framework, offering comprehensive insights into its detection process and optimization methodology. In Section 4, we validate the efficacy of our proposed method through extensive experiments. Lastly, in Section 5, we summarize our contributions and discuss potential avenues for future research.

2. Symbols and Preliminaries

In this section, we introduce the basic notations and give detailed definitions related to the ILN-SSR framework. We use non-bold lowercase letters for scalars, e.g., x, non-bold uppercase letters for two-dimensional matrices, e.g., X, and bold calligraphy letters for three-dimensional tensors, e.g., $\mathcal{X}$.

And for convenience, we use $\mathcal{X}^{(k)}$ to denote the kth frontal slice of tensor $\mathcal{X}$. For a third-order tensor, $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we use MATLAB notations $\mathcal{X}(i,:,:)$, $\mathcal{X}(:,j,:)$, and $\mathcal{X}(:,:,k)$ to denote the horizontal, lateral, and frontal slices, respectively, and we use $\mathcal{X}(i,j,:)$ to represent the tensor tubes, which are vectors along the third dimension of the tensor. We also introduce three block-based operators, bcirc(·), bvec(·), and bvfold(·), defined by (1) and (2):

$$\text{bcirc}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} & \mathcal{X}^{(n_3)} & \dots & \mathcal{X}^{(2)} \\ \mathcal{X}^{(2)} & \mathcal{X}^{(1)} & \dots & \mathcal{X}^{(3)} \\ \vdots & \vdots & \vdots & \vdots \\ \mathcal{X}^{(n_3)} & \mathcal{X}^{(n_3-1)} & \dots & \mathcal{X}^{(1)} \end{bmatrix}, \tag{1}$$

$$\text{bvec}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} \\ \mathcal{X}^{(2)} \\ \vdots \\ \mathcal{X}^{(n_3)} \end{bmatrix}, \quad \text{bvfold}(\text{bvec}(\mathcal{X})) = \mathcal{X}. \tag{2}$$

Definition 1 (t-Product [45]). *Given $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, and Tensor Product $\mathcal{X} * \mathcal{Y}$ can be defined by (3):*

$$\mathcal{X} * \mathcal{Y} = \text{bvfold}(\text{bcirc}(\mathcal{X})\,\text{bvec}(\mathcal{Y})). \tag{3}$$

Theorem 1 (Tensor Singular Value Decomposition (T-SVD) [46]). *Given $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, then it can be factored as (4):*

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*, \tag{4}$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$,$\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal, and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor (meaning each frontal slice is a diagonal matrix with zeros outside the main diagonal). $\mathcal{V}^*$ represents the conjugate transpose of tensor $\mathcal{V}$.

Definition 2 (Tensor Tubal Rank [45,47]). *The tensor tubal rank, denoted as $\text{rank}_t(\cdot)$, is defined as the number of non-zero singular tubes of tensor $\mathcal{S}$, where $\mathcal{S}$ is from the T-SVD of $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$. That is, as shown in Equation (5):*

$$\text{rank}_t(\mathcal{X}) = |\{i : \mathcal{S}(i,i,:) \neq \mathbf{0}\}|. \tag{5}$$

Definition 3 (Logarithmic Norm). *For any matrix $X \in \mathcal{X}^{n_1 \times n_2}$, its matrix logarithmic norm with $\epsilon > 0$ is defined in (6):*

$$\|X\|_L := \sum_{i=1}^{\min\{n_1,n_2\}} \log(\sigma_i(X) + \epsilon), \tag{6}$$

where $\sigma_i(X)$denotes the ith singular value of X.

Theorem 2 (Logarithmic Singular Value Thresholding [42]). *Let $P \in \mathbb{R}^{n_1 \times n_2}$ be a given matrix whose singular value decomposition (SVD) is $P = U_P \Sigma_P V_P^T$, with $\Sigma_P = \text{diag}\Big(\sigma_1(P), \cdots, \sigma_{\min\{n_1,n_2\}}(P)\Big)$, where "diag(...)" represents constructing a diagonal matrix from the input elements in sequence. For any penalty parameter $\alpha > 0$, the global optimal solution to the problem is expressed as (7):*

$$\min_X \frac{1}{2}\|X - P\|_F^2 + \alpha\|X\|_L. \tag{7}$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm [48]. The solution to (7) is $X = U_{\text{p}} \mathbb{L}_{\alpha,\epsilon}(\mathbf{\Sigma p}) V_{\text{P}}^T$, where ϵ is a small positive number used to ensure numerical stability during the logarithmic operation when the singular values approach zero. The operator $\mathbb{L}_{\alpha,\epsilon}$ represents the logarithmic singular value thresholding (LSVT) operator, which is applied element-wise as follows:

$$\mathbb{L}_{\alpha,\epsilon}(x) = \begin{cases} 0, & \Delta \leq 0 \\ \underset{y \in \left\{0, \frac{1}{2}(x-\epsilon+\sqrt{\Delta})\right\}}{\arg\min} h(y), & \Delta > 0 \end{cases}, \tag{8}$$

where $\Delta = (x-\epsilon)^2 - 4(\alpha - x\epsilon)$ *and the function* $h(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ *is defined as*

$$h(y) = \frac{1}{2}(y-x)^2 + \alpha \log(y+\epsilon). \tag{9}$$

3. Materials and Methods

In this section, we present the ILN-SSR framework designed for the robust infrared small target detection. Furthermore, we undertake a comprehensive analysis of this framework, employing the alternating direction method of multipliers optimization algorithm.

3.1. Construction of the Spatial-Temporal Tensor

Drawing upon the research presented in [49], it is elucidated that infrared images can be effectively delineated through the utilization of an additive model, which can be articulated as follows:

$$D(x,y) = B(x,y) + T(x,y) + N(x,y). \tag{10}$$

where D, B, T, N, respectively, represent the original infrared image, background image, target image, and noise image, and (x,y) denotes the corresponding pixel location.

In this study, we strive to comprehensively exploit both spatial and structural information to reconstruct spatial-temporal tensors, aiming to more effectively demonstrate the background low-rankness and inherent sparsity of infrared images.

Based on the analysis in Figure 1d, the tensor blocks corresponding to the background component demonstrate superior low-rank characteristics compared with the global context. Consequently, we have devised an approach, as illustrated in Figure 2, to transform the image sequence into tensor blocks. In this approach, the rth frame, denoted as I_r, is sequentially partitioned into a series of image patches following an "S"-shaped path that progresses from top to bottom. The patch size and slide step, denoted as ps and ss, are specifically designed to ensure that there is no overlap between the generated image patches. The patches are then stacked in sequence, forming a tensor denoted as $\mathcal{D}$, just like Figure 2a shows. To leverage the temporal information, we further stack tensors from p adjacent frames together, as illustrated in Figure 2b. Therefore, by incorporating the representation provided in (10), tensor $\mathcal{D}$ can be decomposed into (11):

$$\mathcal{D}^{st} = \mathcal{B}^{st} + \mathcal{T}^{st} + \mathcal{N}^{st}, \tag{11}$$

here, $\mathcal{D}^{st}$, $\mathcal{B}^{st}$, $\mathcal{T}^{st}$, and $\mathcal{N}^{st}$ represent the spatial-temporal tensors illustrated in Figure 2, corresponding, respectively, to the original image, background, target, and noise.

Since the low-rank property of the background and the sparsity of the target remain valid within the tensor, target detection can be achieved by optimizing the following framework, described as (12):

$$\begin{aligned} &\arg\min_{\mathcal{B},\mathcal{T},\mathcal{N}} \operatorname{rank}(\mathcal{B}) + \lambda_1 \|\mathcal{T}\|_0 + \lambda_2 \|\mathcal{N}\|_F^2 \\ &\text{s.t.} \quad \mathcal{D} = \mathcal{B} + \mathcal{T} + \mathcal{N}, \end{aligned} \tag{12}$$

where the "arg min" notation represents the set of variables that minimize the given function, "s.t." stands for "subject to", λ_1, λ_2 are positive optimization parameters, $\|\cdot\|_0$ stands for l_0-norm, and $\|\cdot\|_F$ stands for the Frobenius norm.

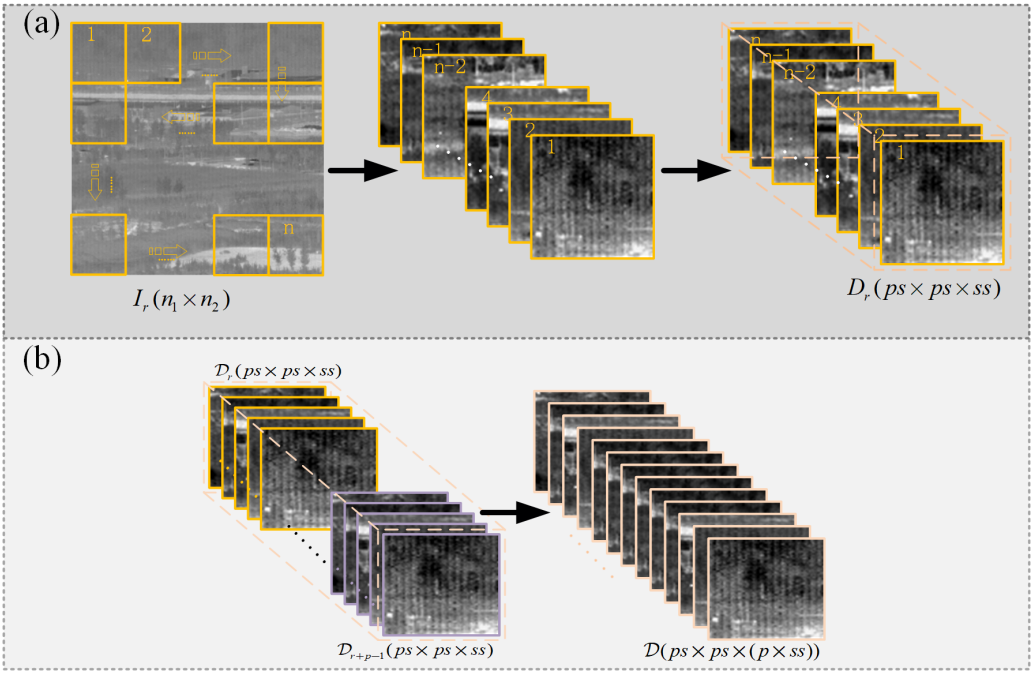


Figure 2. (**a**) Construction of a corresponding tensor from original image, where n_1 and n_2 correspond to the width and height of the constructed image. (**b**) Stack tensors in chronological order.

3.2. Non-Convex Surrogate of Rank

As discussed in Section 1.1, optimizing the background tensor is crucial for IRSTD. However, as noted in [50], directly solving the rank is NP-hard. Introducing a surrogate makes the optimization more feasible while effectively capturing the low-rank nature of the data.

As depicted in Figure 3, the singular values of background image blocks exhibit a long-tailed distribution, with the majority of values clustering near zero and only a few being significantly larger. A simple summation of these singular values tends to disproportionately penalize the larger values while inadequately addressing those near zero. As indicated in Table 1, most methods aimed at rank approximation face challenges in addressing these two aspects concurrently. Inspired by the research presented in [51], we recognized that the logarithmic function, with its steep slope near zero, effectively penalizes small singular values while offering a balanced treatment of larger ones. This ensures that both small and large singular values are appropriately considered in the surrogate method. Furthermore, drawing from the findings in [43,44], we observed that the surrogate method proposed in [42], referred to as the improved logarithmic norm, imposes less penalization on larger singular values and more on smaller ones compared with the logarithmic norm.

Based on the above analysis, we introduce the improved logarithmic norm to the field of infrared small target detection. The improved logarithmic norm can thus be defined as (13):

$$\|\mathcal{X}\|_{ILN} = p \sum_{k=1}^{n_3} \sum_{i=1}^{\min\{n_1,n_2\}} \log\left(\sigma_i^{1/p}(X_k) + \epsilon\right). \tag{13}$$

When $p > 1$, the $\|\mathcal{X}\|_{ILN}$ norm imposes less penalization on larger singular values and more penalization on smaller ones compared with the case when $p = 1$. Additionally, as indicated by (14), the $\|\mathcal{X}\|_{ILN}$ norm exhibits a tighter upper bound [42]:

$$\|\mathcal{X}\|_{ILN} = \min \sum_{i=1}^{p} \|\mathcal{X}_i\|_L^1 \quad \text{s.t.} \quad \mathcal{X} = \mathcal{X}_1 * \mathcal{X}_2 * \ldots * \mathcal{X}_p. \tag{14}$$

According to Zhou et al. [52], decomposing a large-scale tensor into two smaller tensors based on their approach facilitates efficient low-rank estimation. Therefore, the optimization problem $\arg\min \|\mathcal{X}\|_{LN} + f(\mathcal{X})$ can be reformulated as $\arg\min \sum_{i=1}^{n} \|\mathcal{X}_i\|_L^1 + f(\mathcal{X}_1 * \ldots * \mathcal{X}_n)$, effectively reducing the large-scale tensor norm to the norms of smaller tensors, thus enabling more efficient computation.

According to the analysis in this section, it is evident that the improved logarithmic norm performs better compared with other substitutes for rank. By adjusting parameters ϵ and p, ILN can adapt to various complex scenarios.

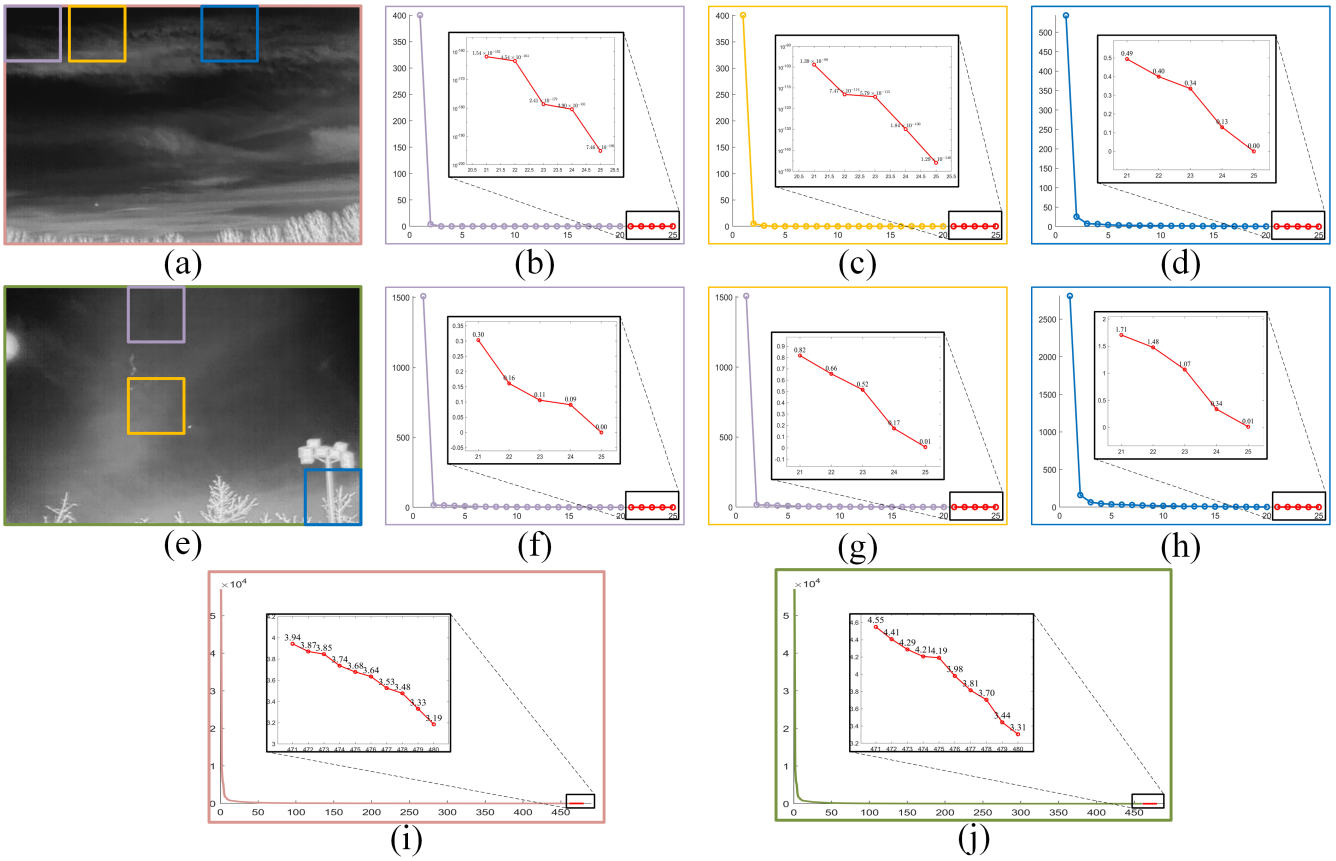


Figure 3. Analysis of low-rank properties. (**b**–**d**) are the singular value trends of the patches in (**a**); (**f**–**h**) are the singular value trends of the patches in (**e**); (**i**,**j**) correspond to the global singular value trends of images (**a**,**e**), respectively.

Table 1. Rank surrogate comparison.

Surrogate Methods	Formula	Balance Large Value	Penalize Small Value
TNN	$\sum_{i=1}^{\min\{n_1,n_2\}} \sigma_i(X)$	no	no
Schatten-p Norm	$\|\mathcal{X}\|_{\mathcal{W},S_p}^p = \frac{1}{L}\sum_{i=1}^{r}\sum_{j=1}^{n_3}(\mathcal{W}(i,i,j)(\overline{\mathcal{S}}(i,i,j))^p)^{\frac{1}{p}}$	yes	no
L_γ Norm	$\|B\|_\gamma = \sum_t \frac{(1+\gamma)\sigma_t(B)}{\gamma+\sigma_t(B)}, \quad \gamma > 0$	yes	no
PSTNN	$\|X\|_{PSTNN} = \sum_{i=p+1}^{\min(n_1,n_2)} \sigma_i(X)$	yes	no
ILN	$\|\mathcal{X}\|_{ILN} = p\sum_{k=1}^{n_3}\sum_{i=1}^{\min\{n_1,n_2\}} \log\left(\sigma_i^{1/p}(X_k) + \epsilon\right)$	yes	yes

3.3. Sparse Structure Refinement

The eight regions in the image that are sparse and distinct from the surrounding background often containing clutter. Consequently, equating small targets with sparse components poses the risk of generating false alarms, thereby compromising detection accuracy.

This section focuses on refining the sparse components by isolating non-target elements from the sparse structure, which we refer to as sparse structure refinement (SSR).

3.3.1. Linear Sparse Regularization Term

In practical scenarios, various components display linearly structured configurations, including isolated roadways, meandering rivers, and flagpoles situated within public squares, as illustrated in Figure 1e,f. These components exhibit a linearly organized sparsity within the complex entirety of infrared imagery. Our research indicates that they constitute a notable source of false alarms.

In order to mitigate the misclassification of such linearly sparse structures as potential targets, we have incorporated the $l_{1,1,2}$-norm in [38] as a reference point. Additionally, we introduced a regularization term for tubewise sparse tensors, based on the l_2-norm of tensor tubes within the optimization framework. This regularization term effectively characterizes linearly sparse arrangements in both the horizontal and vertical directions, proficiently suppressing the emergence of these structures within the target tensor. The formulation of this regularization, based on the l_2-norm of third-order tensor tubes, is delineated as (15):

$$\|\mathcal{E}\|_{ls} = \sum_{i,l} \|\mathcal{E}_1(i,:,l)\|_2 + \sum_{j,l} \|\mathcal{E}_2(:,j,l)\|_2. \tag{15}$$

Here, $\|\cdot\|_2$ represents the l_2-norm, and $\|\cdot\|_{ls}$ denotes the regularization term used to describe linearly structured sparsity. The terms $\mathcal{E}_1$ and $\mathcal{E}_2$ represent the linearly sparse components in the horizontal and vertical directions, respectively.

3.3.2. Target-Traits-Based Sparse Regularization Term

To address the nonlinear sparse components in the image, which contribute to false alarms due to their point-like, block-like, or irregular shapes not aligning with the regularization terms in Section 3.3.1, this study introduces an optimization term that integrates target feature descriptions into the framework to enhance sparse structure refinement.

In the context of infrared imagery, it is worth noting that small targets often exhibit distinct point spread characteristics, wherein the majority of local gradients converge toward their central points. Upon applying Gaussian filtering, these targets experience enhancement, while elements within the background that do not adhere to these specific characteristics are subdued.

In response to this, we introduce an innovative regularization term, sensitized to the distinctive traits of targets, represented as the l_1-norm resulting from the convolution between a structural tensor, $\mathcal{P}$, and a filter kernel, K, which is chosen based on target characteristics. This sparse optimization term filters out components in the tensor $\mathcal{D}$ that do not match expected target features, efficiently suppressing unwanted sparse elements. In light of the aforementioned exposition, the target feature-based sparse regularization term is formally defined as (16):

$$\|\mathcal{P}\|_{KF} = \|K \circledast \mathcal{P}\|_1, \tag{16}$$

where $\|\cdot\|_1$ represents the l_1-norm, $\|\cdot\|_{KF}$ stands for the sparse optimization term based on target traits, and $\circledast$ represents the convolution operation.

The filter kernel, K, can be obtained through visual estimation or by performing simple calculations based on the target. Figure 4 illustrates the formation process of the filtering kernel, K. Initially, infrared point targets are downsampled to dimensions of 5×5 and

subsequently quantized to integer values ranging from 0 to 8. This process can be succinctly described as (17):

$$K = \lfloor \text{Extract}(\mathcal{I}_t\}) / Interval \rfloor, \tag{17}$$

where $\mathcal{I}_t$ is the target image, and $Interval$ represents the quantization interval, which, for an 8-bit grayscale image quantized to values between 0 and 8, equals 32.

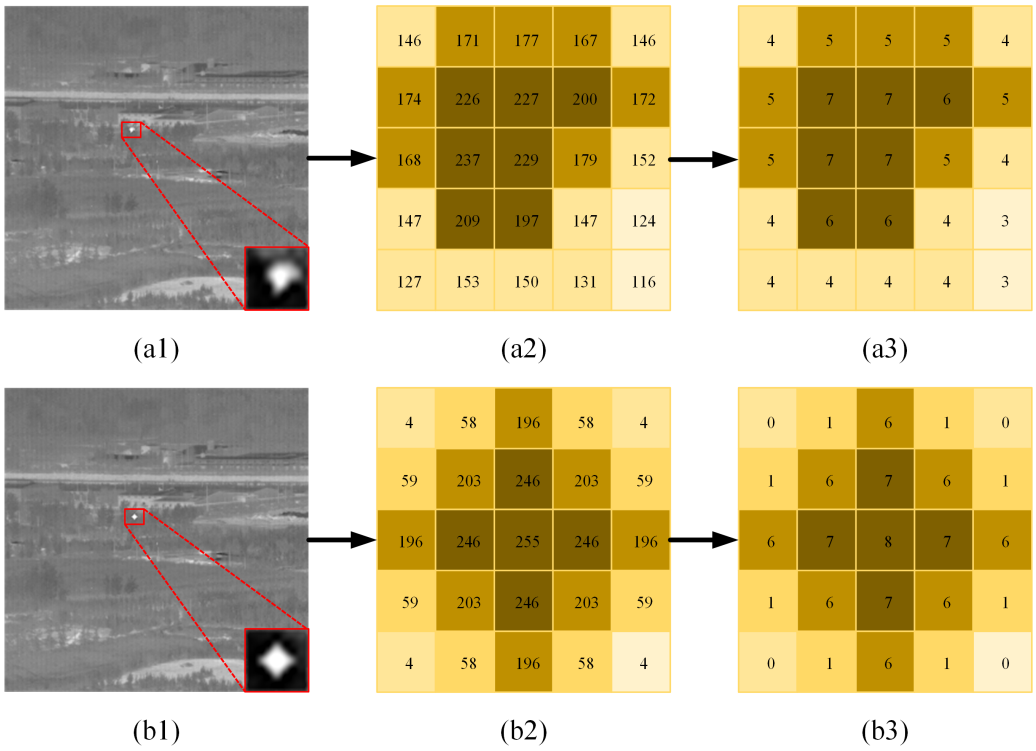


Figure 4. Formation of kernel $\mathcal{K}$, where (**a1**,**b1**) represents the IR image with highlighted targets, (**a2**,**b2**) is extracted target 5×5 neighborhood, and (**a3**,**b3**) are the generated result, $\mathcal{K}$.

3.4. Local Prior Weight Map

The use of prior weights can effectively enhance LRSD-based object detection. In essence, the prior weight acts as a precursor to detection, accentuating regions of interest and expediting optimization convergence. To optimize performance, the prior weight should emphasize absent image textures and local statistical properties within the optimization framework. Here, we introduce the concept of image entropy in (18):

$$E(x, y) = -\sum_{i=0}^{m} (f_i - f_{\text{center}})^2 \log(p(i)), \tag{18}$$

where m denotes the quantity of grayscale levels encompassed within the image neighborhood, each corresponding to pixel grayscale values denoted as $f_i (i = 1, 2, \ldots, m)$, and their corresponding probability distribution as $p_i (i = 1, 2, \ldots, m)$.

In the research conducted by Gao et al. [53], it was empirically established that structural tensors possess the capacity to efficiently encapsulate the structural characteristics of an image, all while incurring a modest computational burden. In accordance with this finding, the present study employs the computation of the structural tensor predicated on the outcomes derived from Equation (18). The approach for calculating the structural tensor, utilizing the outcomes of information entropy, is expounded upon as (19):

$$ST = \begin{bmatrix} G_\sigma \circledast g_x^2 & G_\sigma \circledast g_x g_y \\ G_\sigma \circledast g_x g_y & G_\sigma \circledast g_y^2 \end{bmatrix}, \tag{19}$$

where G_σ denotes the Gaussian kernel with a variance denoted as sigma, and $g_x = \partial E/\partial x$ and $g_y = \partial E/\partial y$ represent the derivatives in two directions.

Each point within the image is associated with a 2×2 structure tensor, which is characterized by a pair of eigenvalues, λ_1 and λ_2. These eigenvalues serve as dual characteristic descriptors of the local geometric structures. Specifically, in flat regions, $\lambda_1 \approx \lambda_2 \approx 0$; in edge regions, $\lambda_1 \gg \lambda_2 \approx 0$; and, in corner regions, $\lambda_1 \geq \lambda_2 \gg 0$. Considering these characteristics, a target-sensitive factor, denoted as w_{ts} in (20), is introduced to emphasize regions within the image that exhibit corner features:

$$w_{ts}(x,y) = \frac{\det(ST(x,y))}{\mathrm{tr}(ST(x,y))} = \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}. \tag{20}$$

Additionally, the factor w_{bs} in (21) demonstrates sensitivity to both edges and targets, for the purpose of extracting edge-related information:

$$w_{bs}(x,y) = \max(\lambda_1, \lambda_2). \tag{21}$$

This research integrates the two aforementioned factors and subsequently performs normalization to derive a local prior feature map, and, to couple with the target-traits-based sparse regularization term, we use the filter mentioned before to expand the highlighted area, as in (22):

$$W_p = \mathrm{Normalize}(w_{ts} \cdot w_{bs}) \circledast K, \tag{22}$$

and to facilitate the identification of candidate regions suitable for inclusion in the optimization framework, the local prior weight map, W_p, is transformed into a local prior weight tensor, $\mathcal{W}_p$, using the method described in Section 3.1. It is worth noting that, functioning as a prior weight, this transformation enhances the optimization process by accelerating the rate of convergence. It is evident that the local prior weight introduced in this section differentially emphasizes potential target regions while simultaneously mitigating the influence of extraneous elements.

3.5. The Proposed Optimization Framework

In contrast to the decomposition presented in Equation (12), the optimization framework developed in this study focuses on both the refinement of sparse structural components and the use of the improved logarithmic norm to estimate low-rank properties. Within the framework outlined earlier in this paper, the tensor $\mathcal{D}$ can be decomposed into several constituent tensors: the background tensor $\mathcal{B}$, the target tensor $\mathcal{T}$, the tensor $\mathcal{E}_1$ representing column sparsity, the tensor $\mathcal{E}_2$ representing row sparsity, the tensor $\mathcal{P}$ delineating sparse structures related to non-target attributes, and the tensor $\mathcal{N}$ representing noise.

After substituting the rank with ILN, we integrated the regularization terms elucidated in Section 3.3 to facilitate a more sophisticated refinement of sparse structural components, alongside the introduction of a discrete noise term. Ultimately, our proposed framework for the detection of infrared small targets can be succinctly described as (23):

$$\begin{array}{l} \min \|\mathcal{B}\|_{ILN} + \lambda_1\|\mathcal{W} \odot \mathcal{T}\|_1 + \lambda_2\|\mathcal{N}\|_F + \lambda_3\|\mathcal{E}\|_{ls} + \lambda_4\|\mathcal{P}\|_{KF} \\ \text{s.t. } \mathcal{D} = \mathcal{B} + \mathcal{T} + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{P} + \mathcal{N} \end{array}, \tag{23}$$

where $\odot$ stands for element-wise product between tensors, and λ_1, λ_2, λ_3, λ_4 stand for positive tradeoff parameters.

In the research conducted by Candès et al. [54], the incorporation of weight factors is specifically aimed at mitigating the influence of signal magnitude on the l_1 penalty function. Here, weight factors describe the factors representing the weight of each point in the target tensor, which are negatively correlated with the values of the target. Utilizing the reciprocal of the signal magnitude as weights in subsequent iterations proves to be

an effective strategy. By iteratively applying the results from the previous iteration, the required weight factors can be obtained as (24):

$$\mathcal{W}_{\text{optim}}^{p+1} = \frac{a}{|\mathcal{T}^p| + b}. \tag{24}$$

The coefficient a is adjustable and set to 2 in the experiment, b is a small positive constant specifically introduced to circumvent division by zero scenarios, and p is utilized to signify the count of iterations.

By combining the prior local tensor, as explained in Section 3.4, with the results from (24), we create weight tensors denote as $\mathcal{W}^k$, specifically designed for the l_1 penalty function used in this (25):

$$\mathcal{W}^k = \mathcal{W}_{pr} \odot \mathcal{W}_{\text{optim}}^k, \tag{25}$$

where $\mathcal{W}_{pr}$ represents the element-wise reciprocal of $\mathcal{W}_p$, and k is utilized to signify the count of iterations.

3.6. ADMM Solver for ILN-SSR

The optimization problem (23) can be solved using ADMM, which is a powerful optimization technique that decomposes a complex problem into smaller subproblems, each of which can be solved more easily. ADMM is particularly effective in problems where variables can be optimized alternately and separately, as it enables the problem to converge efficiently by iterating over subproblems. By introducing auxiliary parameters $\mathcal{Z}$ and $\mathcal{V}$, we can obtain an equivalent expression for (23) as (26):

$$\begin{array}{l} \min \|\mathcal{Z}\|_{ILN} + \lambda_1 \|\mathcal{W} \odot \mathcal{T}\|_1 + \lambda_2 \|\mathcal{N}\|_F + \lambda_3 \|\mathcal{E}\|_{Is} + \lambda_4 \|\mathcal{V}\|_1 \\ \text{s.t. } \mathcal{D} = \mathcal{B} + \mathcal{T} + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{P} + \mathcal{N}, \mathcal{Z} = \mathcal{B}, \mathcal{V} = K \circledast \mathcal{P} \end{array}. \tag{26}$$

Here, $\mathcal{B}$ represents the background tensor, $\mathcal{T}$ represents the target tensor, $\mathcal{W}$ represents the weight tensor, $\mathcal{N}$ represents the noise tensor, $\mathcal{E}_1, \mathcal{E}_2$ represent the linearly sparse components, $\mathcal{P}$ represents the non-target sparse components, and $\mathcal{Z}$ and $\mathcal{K}$ represent auxiliary components used for optimization. The augmented Lagrangian function of (26) is given as follows:

$$\begin{array}{l} \mathcal{L}(\mathcal{B}, \mathcal{T}, \mathcal{E}_1, \mathcal{E}_2, \mathcal{P}, \mathcal{N}, \mathcal{Z}, \mathcal{V}) \\ = \|\mathcal{Z}\|_{ILN} + \lambda_1 \|\mathcal{W} \odot \mathcal{T}\|_1 + \lambda_2 \|\mathcal{N}\|_F + \lambda_3 \|\mathcal{E}\|_{Is} + \lambda_4 \|\mathcal{V}\|_1 \\ + \langle y_1, \mathcal{Z} - \mathcal{B} \rangle + \langle y_2, \mathcal{D} - \mathcal{B} - \mathcal{T} - \mathcal{E}_1 - \mathcal{E}_2 - \mathcal{P} - \mathcal{N} \rangle \\ + \langle y_3, \mathcal{V} - K \circledast \mathcal{P} \rangle + \frac{\mu}{2} \left(\|\mathcal{Z} - \mathcal{B}\|_F^2 \right. \\ \left. + \|\mathcal{D} - \mathcal{B} - \mathcal{T} - \mathcal{E}_1 - \mathcal{E}_2 - \mathcal{P} - \mathcal{N}\|_F^2 + \|\mathcal{V} - K \circledast \mathcal{P}\|_F^2 \right), \end{array} \tag{27}$$

where μ represents a positive penalty scalar and y_1, y_2, y_3 represent the Lagrangian multipliers. The ADMM algorithm decomposes (27) into several optimization subproblems. Since it is difficult to optimize all the variables in (27) simultaneously, the variables are updated alternately. For instance, when optimizing $\mathcal{B}$, the subproblem only considers the terms related to $\mathcal{B}$.

3.6.1. Subproblem of $\mathcal{Z}$

In accordance with the formalized definition of the improved logarithmic norm, as elucidated in Section 3.2, it becomes apparent that, by expanding the variable $\mathcal{Z}$, we

can articulate the subproblem pertaining to the auxiliary variable $\mathcal{Z}$ in a manner that establishes equivalence.

$$\begin{aligned}\mathcal{Z}^{k+1} &= \arg\min_{\mathcal{Z}} \|\mathcal{Z}\|_{ILN} + \frac{\mu^k}{2}\left\|\mathcal{Z} - \mathcal{B}^k + \frac{y_1^k}{\mu^k}\right\|_F^2 \\ &= \arg\min_{Z} \sum_{i=1}^{n} \|\mathcal{Z}_i\|_L^1 + \frac{\mu^k}{2}\left\|\mathcal{Z}_1 * \mathcal{Z}_2 \ldots * \mathcal{Z}_n - \mathcal{B}^k + \frac{y_1^k}{\mu^k}\right\|_F^2,\end{aligned} \tag{28}$$

where k represents the number of iterations, and the sub-optimization problem for the ith factor $\mathcal{Z}_i$ in (28) can be described as

$$\mathcal{Z}_i^{k+1} = \arg\min_{\mathcal{Z}_i} \|\mathcal{Z}_i\|_L^1 + \frac{\mu^k}{2}\left\|\mathcal{Q}_{1,i}^k * \mathcal{Z}_i * \mathcal{Q}_{2,i}^k - \mathcal{B}^k + \frac{y_1^k}{\mu^k}\right\|_F^2, \tag{29}$$

where $\mathcal{Q}_{1,i}^k = \mathcal{Z}_1^k * \mathcal{Z}_2^k * \ldots * \mathcal{Z}_{i-1}^k$ and $\mathcal{Q}_{2,i}^k = \mathcal{Z}_{i+1}^k * \mathcal{Z}_{i+2}^k * \ldots * \mathcal{Z}_n^k$ are used for simplicity. In relation to Equation (29), a Taylor series expansion can be applied to the quadratic term, centered around a specified extrapolation point. Following this methodological approach, the equation is subsequently transformed into (30):

$$\mathcal{Z}_i^{k+1} = \arg\min_{\mathcal{Z}_i} \|\mathcal{Z}_i\|_L^1 + \frac{\xi_i^k}{2}\left\|\mathcal{Z}_i - \left[\mathcal{Z}_i^k\right] + \frac{1}{\xi_i^k} d_i\left(\left[\mathcal{Z}_i^k\right]\right)\right\|_F^2, \tag{30}$$

where ξ_i^k is a proximal parameter and $\left[\mathcal{Z}_i^k\right]$ is the extrapolation point, $d_i\left(\left[\mathcal{Z}_i^k\right]\right) = \mu^k \cdot \mathcal{Q}_{1,i}^{k,T} * \left(\mathcal{B}^k - \frac{y_1^k}{\mu^k} - \mathcal{Q}_{1,i}^k * \mathcal{Z}_i^k * \mathcal{Q}_{2,i}^k\right) * \mathcal{Q}_{2,i}^{k,T}$. By using LSVT in Section 2, we can update $\mathcal{Z}$ using the (31):

$$\overline{\mathcal{Z}}_i^{k+1,(j)} = \bar{U}_{\mathcal{M}}^{(j)} \mathbb{L}_{1/\xi_i^k,\epsilon}\left(\bar{\Sigma}_{\mathcal{M}}^{(j)}\right) \bar{V}_{\mathcal{M}}^{(j),T}, \tag{31}$$

where $\mathcal{M}^{(j)} = \bar{U}_{\mathcal{M}}^{(j)} \bar{\Sigma}_{\mathcal{M}}^{(j)} \bar{V}_{\mathcal{M}}^{(j),T}$ denotes the jth frontal slice of the DFT of $\mathcal{M} = \left[\mathcal{Z}_i^k\right] - \frac{1}{\xi_i^k} d_i\left(\left[\mathcal{Z}_i^k\right]\right)$ along the third dimension, and $\mathbb{L}_{a,\epsilon}(\cdot)$ is defined in Section 2 Theorem 2.

And we can efficiently derive the closed-form solution of Equation (29): $\mathcal{Z}_i^{k+1} = \mathcal{F}^{-1}\left(\overline{\mathcal{Z}}_i^{k+1}\right)$, where $\mathcal{F}^{-1}(\cdot)$ represents the inverse discrete Fourier transform, and $\overline{\mathcal{Z}}_i^{k+1}$ is a third-order tensor whose jth frontal slice is $\overline{\mathcal{Z}}_i^{k+1,(j)}$.

$\mathcal{Z}$ can be updated via (32):

$$\mathcal{Z}^{k+1} = \mathcal{Z}_1^{k+1} * \mathcal{Z}_2^{k+1} * \ldots * \mathcal{Z}_n^{k+1}. \tag{32}$$

3.6.2. Subproblem of $\mathcal{B}$

After obtaining $\mathcal{Z}$,the solution for $\mathcal{B}$ in (33) can be found by taking the derivative and finding the minimum:

$$\begin{aligned}\mathcal{B}^{k+1} &= \arg\min_{\mathcal{B}} \frac{\mu^k}{2}\left\|\mathcal{Z}^{k+1} - \mathcal{B} + \frac{y_1^k}{\mu^k}\right\|_F^2 + \frac{\mu^k}{2}\left\|\mathcal{D} - \mathcal{B} - \mathcal{T}^k - \mathcal{E}_1^k - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k}\right\|_F^2 \\ &= \frac{\mu^k\left(\mathcal{Z}^{k+1} + \mathcal{D} - \mathcal{T}^k - \mathcal{E}_1^k - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k\right) + y_1^k + y_2^k}{2\mu^k}.\end{aligned} \tag{33}$$

3.6.3. Subproblem of $\mathcal{T}$

The solution of $\mathcal{T}$ can be described as (34):

$$\begin{aligned}\mathcal{T}^{k+1} &= \arg\min_{\mathcal{T}} \lambda_1 \left\| \mathcal{W}^k \odot \mathcal{T} \right\|_1 \\ &+ \frac{\mu^k}{2} \left\| \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T} - \mathcal{E}_1^k - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k} \right\|_F^2,\end{aligned} \tag{34}$$

and the aforementioned equation can be solved by employing the Iterative Soft Thresholding Algorithm [55], as shown in (35):

$$\mathcal{T}^{k+1} = \mathcal{S}_{\lambda_1 \mathcal{W}^k / \mu^k} \left(\mathcal{D} - \mathcal{B}^{k+1} - \mathcal{E}_1^k - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k} \right), \tag{35}$$

where $\mathcal{S}(\cdot)$ represent the soft threshold shrinkage operator $\mathcal{S}_\xi(x) = \text{sign}(x) \cdot \max(|x| - \xi, 0)$.

3.6.4. Subproblem of $\mathcal{W}$

The weight $\mathcal{W}$ can be updated by combining $\mathcal{T}$ with the local prior weight map, as in (36):

$$\mathcal{W}^{k+1} = \mathcal{W}_{pr} \odot \mathcal{W}_{\text{optim}}^k . \tag{36}$$

3.6.5. Subproblem of $\mathcal{E}_1$,$\mathcal{E}_2$

The $\mathcal{E}_1$ can be updated via (37):

$$\begin{aligned}\mathcal{E}_1^{k+1} &= \arg\min_{\mathcal{E}_1} \lambda_3 \textstyle\sum_{i,l} \|\mathcal{E}_1(i,:,l)\|_2 \\ &+ \frac{\mu^k}{2} \left\| \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{E}_1 - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k} \right\|_F^2.\end{aligned} \tag{37}$$

Let $\mathcal{S} = \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{E}_2^k - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k}$. The closed-form solution to (37) can be obtained based on the conclusions from (38):

$$\mathcal{E}_1^{k+1}(:,c,:) = \left(1 - \frac{\lambda_3}{\mu^k \|\mathcal{S}(:,c,:)\|_2} \right)_+ \mathcal{S}(:,c,:). \tag{38}$$

A similar process can be applied to derive the solution for $\mathcal{E}_2$:

$$\mathcal{E}_2^{k+1}(r,:,:) = \left(1 - \frac{\lambda_3}{\mu^k \|\mathcal{S}'(r,:,:)\|_2} \right)_+ \mathcal{S}'(r,:,:), \tag{39}$$

where $\mathcal{S}' = \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{E}_1^{k+1} - \mathcal{P}^k - \mathcal{N}^k + \frac{y_2^k}{\mu^k}$

3.6.6. Subproblem of $\mathcal{P}$

The subproblem of $\mathcal{P}$ can be described by Equation (40):

$$\begin{aligned}\mathcal{P}^{k+1} &= \arg\min_{\mathcal{P}} \frac{\mu^k}{2} \left\| \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{E}_1^{k+1} - \mathcal{E}_2^{k+1} - \mathcal{P} - \mathcal{N}^k + \frac{y_2^k}{\mu^k} \right\|_F \\ &+ \frac{\mu^k}{2} \left\| \mathcal{Y}^k - K \circledast \mathcal{P} + \frac{y_3^k}{\mu^k} \right\|_F^2.\end{aligned} \tag{40}$$

According to the research content of [56], by altering the spatial arrangement of the convolution kernel and tensor, we can transmute the convolution operation into an equivalent tensor multiplication, thereby transfiguring Equation (40) into (41):

$$\begin{aligned}\left(\mathcal{I}+\mathcal{K}^{\mathrm{T}}*\mathcal{K}\right)*\mathcal{P}^{k+1,v}=&\left(\mathcal{D}-\mathcal{B}^{k+1}-\mathcal{T}^{k+1}-\mathcal{E}_1^{k+1}-\mathcal{E}_2^{k+1}-\mathcal{N}^k+\frac{y_2^k}{\mu^k}\right)\\&+\mathcal{K}^{\mathrm{T}}*\left(\mathcal{V}^{k,v}+\frac{y_3^{k,v}}{\mu^k}\right),\end{aligned}\tag{41}$$

in the presented framework, $\mathcal{I}$ denotes the identity tensor, while $\mathcal{K}$ symbolizes the tensor representative of the filter, which is derived post-convolution and subsequently transformed into an equivalent multiplicative form. We can solve to obtain the solution for $\mathcal{P}$ as in (42):

$$\mathcal{P}^{k+1}=\mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathcal{R})+\mathrm{psf2otf}(K,\mathrm{Size})^H\cdot*\mathcal{F}(\mathcal{X})}{1+\mathrm{psf2otf}(K,\mathrm{Size}).*\mathrm{psf2otf}(K,\mathrm{Size})^H}\right),\tag{42}$$

where $\mathcal{R}=\mathcal{D}-\mathcal{B}^{k+1}-\mathcal{T}^{k+1}-\mathcal{E}_1^{k+1}-\mathcal{E}_2^{k+1}-\mathcal{N}^k+\frac{y_2^k}{\mu^k}$, $\mathcal{X}=\mathcal{V}^k+\frac{y_3^k}{\mu^k}$ are used for simplicity, $\mathcal{F}(\cdot)$ is the nonuniform fast Fourier transform (nFFT) [57], $\mathcal{F}^{-1}(\cdot)$ is the inverse nonuniform fast Fourier transform, and size stands for the size of tensor $\mathcal{D}$. "psf2otf" is a MATLAB built-in function that transforms a spatial input pattern into its frequency domain equivalent by reorganizing its elements and applying the Fourier transform.

3.6.7. Subproblem of $\mathcal{V}$

Variable $\mathcal{V}$ can be updated with Equation (43):

$$\mathcal{V}^{k+1}=\arg\min_{\mathcal{V}}\lambda_4\|\mathcal{V}\|_1+\frac{\mu^k}{2}\left\|\mathcal{V}-K\circledast\mathcal{P}^{k+1}+\frac{y_3^k}{\mu^k}\right\|_F^2\tag{43}$$

Similar to subproblem $\mathcal{T}$, we can use the soft-thresholding operator to derive the solution to the problem, as described in Equation (44).

$$\mathcal{V}^{k+1}=\mathcal{S}_{\lambda_4/\mu^k}\left(K\circledast\mathcal{P}^{k+1}-\frac{y_3^k}{\mu^k}\right)\tag{44}$$

3.6.8. Subproblem of $\mathcal{N}$

The variable $\mathcal{N}$ can be updated by taking the derivative to obtain the extremum (45):

$$\begin{aligned}\mathcal{N}^{k+1}&=\arg\min_{\mathcal{N}}\lambda_2\|\mathcal{N}\|_F^2+\frac{\mu^k}{2}\left\|\mathcal{D}-\mathcal{B}^{k+1}-\mathcal{T}^{k+1}-\mathcal{E}_1^{k+1}-\mathcal{E}_2^{k+1}-\mathcal{P}^{k+1}-\mathcal{N}+\frac{y_2^k}{\mu^k}\right\|_F^2\\&=\frac{\mathcal{D}-\mathcal{B}^{k+1}-\mathcal{T}^{k+1}-\mathcal{E}_1^{k+1}-\mathcal{E}_2^{k+1}-\mathcal{P}^{k+1}+y_2^k}{2\lambda_2+\mu^k}\end{aligned}\tag{45}$$

3.6.9. Subproblem of y_1, y_2, y_3, μ

Based on the citations [40,41], the positive penalty scalar, μ, and the Lagrange multipliers y_1, y_2, y_3 are updated while keeping other variables fixed in (46)–(49):

$$y_1^{k+1}=\mu^k\left(\mathcal{Z}^{k+1}-\mathcal{B}^{k+1}\right)+y_1^k\tag{46}$$

$$y_2^{k+1}=\mu^k\left(\mathcal{D}-\mathcal{T}^{k+1}-\mathcal{N}^{k+1}-\mathcal{B}^{k+1}-\mathcal{E}_1^{k+1}-\mathcal{E}_2^{k+1}-\mathcal{P}^{k+1}\right)+y_2^k\tag{47}$$

$$y_3^{k+1}=\mu^k\left(\mathcal{V}^{k+1}-K\circledast\mathcal{P}^{k+1}\right)+y_3^k\tag{48}$$

$$\mu^{k+1}=\min\left(\kappa\mu^k,\mu_{\max}\right)\tag{49}$$

3.6.10. Iteration Criterion

The iteration process shall terminate and yield the detection result when any of the following conditions are satisfied:

(1) $\frac{\|\mathcal{D}-\mathcal{B}-\mathcal{T}-\mathcal{E}_1-\mathcal{E}_2-\mathcal{P}-\mathcal{N}\|_F}{\|\mathcal{D}\|_F} < \delta$.
(2) When the count of non-zero elements in the target tensor $\mathcal{T}$ remains constant.
(3) Reaching the maximum number of iterations.

The detailed algorithm for addressing the optimization framework using the alternating direction method of multipliers is systematically outlined in Algorithm 1. It iteratively optimizes the background, target, and various sparse structure tensors based on the equations from (30)–(49). Once the termination conditions mentioned in Section 3.6.10 are satisfied, the target-background tensor is reconstructed and output. Correspondingly, Figure 5 provides a more intuitive depiction of the iterative process, clearly illustrating the roles of iteration, construction, and weight tensors in the solving algorithm, as well as the optimization process. In summary, this framework facilitates the decomposition of tensor $\mathcal{D}$ into the background tensor $\mathcal{B}$, the target tensor $\mathcal{T}$, and additional components. This strategic decomposition is pivotal in proficiently discerning targets amidst multifaceted backgrounds and pervasive noise.

Algorithm 1: Optimization of the ILN-SSR Framework

Input: Infrared image sequence $I_r(r = 1, 2, \ldots, k) \in \mathbb{R}^{n_1 \times n_2}$
Output: Target image sequence $T_r(r = 1, 2, \ldots, k) \in \mathbb{R}^{n_1 \times n_2}$, Background image sequence $B_r(r = 1, 2, \ldots, k) \in \mathbb{R}^{n_1 \times n_2}$

Construct image sequence into tensor $\mathcal{D}$ via Section 3.1.
Calculate $\mathcal{W}_p$ via Section 3.4.
Calculate the reciprocal of each term in $\mathcal{W}_p$, denoted as $\mathcal{W}_{pr}$.
Initialize parameters: $\mathcal{B}^0 = \mathcal{P}^0 = \mathcal{D}$, $\mathcal{T}^0 = \mathcal{E}_1^0 = \mathcal{E}_2^0 = \mathcal{V}^0 = \mathcal{N}^0 = 0$, $y_1^0 = y_2^0 = y_3^0 = 0, \mu^0 = 20, \mu_{\max} = 1e5, \kappa = 1.5, a = 2, b = 1$, max_iter = 400, $k = 0$.
while *not converged* **do**
    Update $\mathcal{Z}^{k+1}$ by (30)–(32);
    Update $\mathcal{B}^{k+1}$ by (33);
    Update $\mathcal{T}^{k+1}$ by (35);
    Update $\mathcal{W}^{k+1}$ by (36);
    Update $\mathcal{E}_1^{k+1}$ by (38);
    Update $\mathcal{E}_2^{k+1}$ by (39);
    Update $\mathcal{P}^{k+1}$ by (42);
    Update $\mathcal{V}^{k+1}$ by (44);
    Update $\mathcal{N}^{k+1}$ by (45);
    Update y_1, y_2, y_3, μ by (46)–(49);
    $k = k + 1$;
end
Reconstruct tensor $\mathcal{T}_i$,$\mathcal{B}_i$, into target sequence $T_i(i = 1, 2, \ldots, p)$, and background sequence $B_i(i = 1, 2, \ldots, p)$.

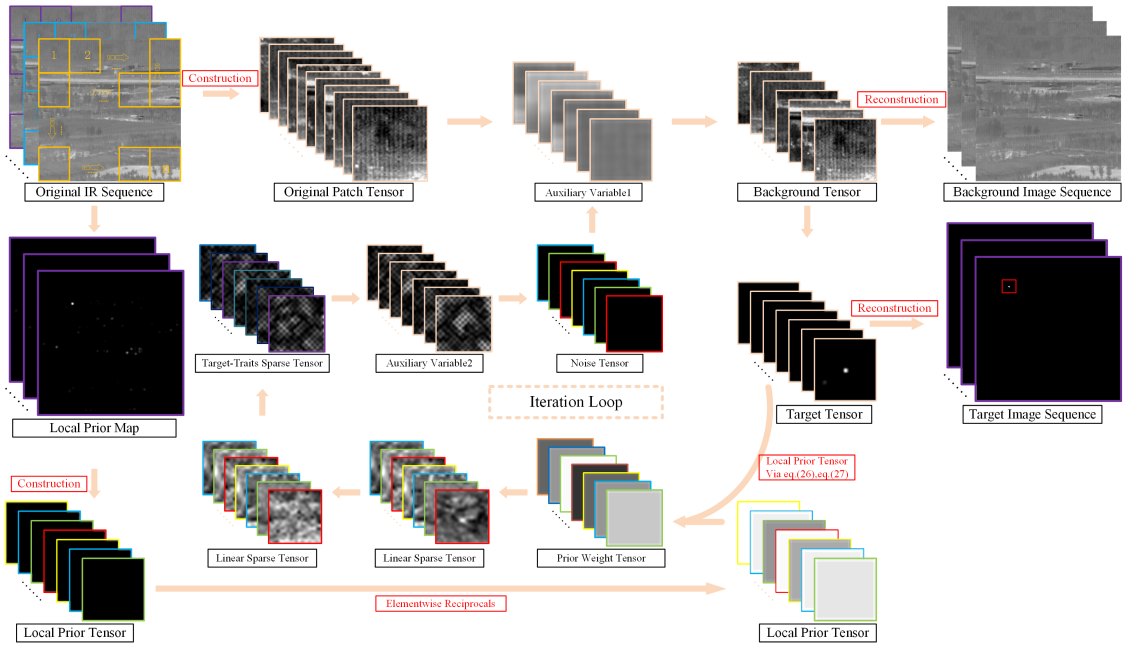


Figure 5. Overview of the ILN-SSR framework.

4. Experiments and Analysis

In this section, a comparative analysis is conducted between the proposed method and a range of state-of-the-art (SOTA) methods.

This section begins by providing a detailed description of the experiment, including the selection and characterization of the dataset, an overview of the comparative methods, and an in-depth introduction to the evaluation criteria. Next, it presents an analysis of the parameters employed in the optimization framework. This is followed by a comprehensive experimental comparison, where the proposed method is evaluated against eight contemporary SOTA methods. To assess the effectiveness of each module within our framework, a series of ablation experiments are conducted. Finally, the section concludes with a discussion of running time analysis.

4.1. Experimental Setup

4.1.1. Datasets

To assess the robustness of our algorithm, we used the SIRST dataset [57] from the National University of Defense Technology and the IRST dataset [58] from Beihang University, selecting six image sequences that feature diverse scenarios and varying sizes.

Figure 6 presents the selected frames from these sequences, highlighting their three-dimensional grayscale distributions, where the horizontal and vertical axes represent the image coordinates, and the third axis represents the grayscale value at each point. The sequences, labeled as Seq.1 to Seq.6 from left to right, have target areas distinctly marked and magnified for clarity. The dataset includes a range of scenarios—multiple targets, point-like and linear clutter, complex backgrounds, and dim targets against simpler backgrounds—offering a comprehensive assessment of the algorithm's robustness.

4.1.2. Baseline Methods

For the purpose of comparative analysis, several infrared small target detection algorithms were employed as benchmarks.

These included detection methods based on the human visual system, such as TLLCM [19] and DNGM [21], the FKRW algorithm derived from the random walk model [10], and the deep learning-based UIUNet [25]. LRSD-based single-frame detection methods, like IPI [31] and NRAM [39], along with multi-frame algorithms such as NTLA [36] and IMNN-LWEC [40], were also employed.

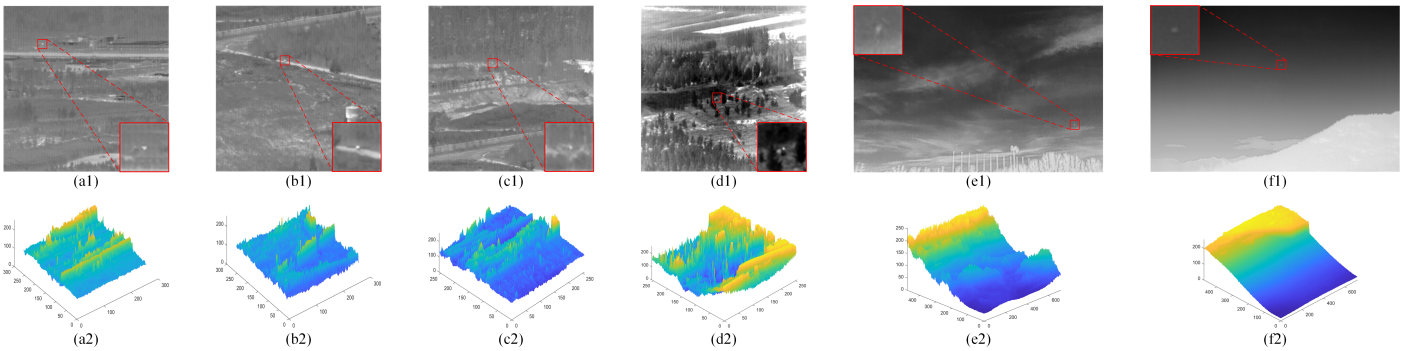


Figure 6. Images and 3-D surface diagrams of six sequences: (**a1**–**f1**) display infrared images from Seq.1 to Seq.6, with targets highlighted by red boxes and enlarged in the image corners, and (**a2**–**f2**) depict the corresponding 3-D surface mesh diagrams for each sequence.

The specific parameter settings for each algorithm are detailed in Table 2. All experiments were conducted on an Ubuntu 20.04 system with an Intel Core i9-13900K CPU @ 5.8 GHz, 64 GB of RAM, and a NVIDIA RTX 4090 GPU with 24 GB of memory. Deep learning-based methods were implemented using Python, while the other methods were implemented in MATLAB.

Table 2. Detailed parameter settings for evaluated methods.

Method	Parameters
FKRW [10]	K = 4, p = 6, β = 200, Window Size: 11 × 11.
TLLCM [19]	K = 9, Gaussian filter size: 3 × 3
DNGM [21]	First Neighbor Size: 7 × 7, Second Neighbor Size: 13 × 13
IPI [31]	Patch Size: 50 × 50, Sliding Step: 10 × 10, $\lambda = 1/\sqrt{\min(m,n)}$, $\varepsilon = 10^{-4}$.
NTLA [36]	L = 3, H = 10, $\lambda_1 = 0.005, \lambda_2 = H/\sqrt{\max(m,n) \times L}, \lambda_3 = 100$
NRAM [39]	Patch Size: 50 × 50, Sliding Step: 10 × 10, $\lambda = 1/\sqrt{\min(n_1,n_2)}$, $\varepsilon = 10^{-7}, \mu_0 = 3\sqrt{\min(n_1,n_2)}, \gamma = 0.002, C = \sqrt{\min(n_1,n_2)}/2.5$
IMNN-LWEC [40]	Frame: 3, Patch Size: 15 × 15, Sliding Step: 15 × 15, $\omega = 5, \lambda_1 = \lambda_L/\sqrt{\max(n_1,n_2) \times n_3}$, $\lambda_2 = \lambda_3 = 50\lambda_1, \lambda_L = 1.4, \beta_{kk_2} = \rho = 1/\omega, \kappa = 1.5,\ \varepsilon = 10^{-2}$
Proposed	Patch Size = 25, Sliding Step = 25, p = 3, $\mu_0 = 20; \lambda_1 = 1.4/\sqrt{\max(n_1,n_2) \times n_3}, \delta = 10^{-4}$

4.1.3. Evaluation Metrics

To rigorously assess the efficacy of the algorithm, this study adopts a multifaceted evaluation approach. This approach incorporates the signal-to-clutter ratio gain (SCRG) and the Background Suppression Factor (BSF), along with an analysis of the Receiver Operating Characteristic (ROC) curve. These metrics collectively evaluate the small target detection algorithm across three key dimensions: target visibility enhancement, background suppression effectiveness, and overall detection capability.

The BSF is a metric dedicated to evaluating the algorithm's proficiency in attenuating background elements. It is quantitatively defined as the ratio of the standard deviation of the processed image to that of the original image:

$$BSF = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}} + \varepsilon}, \tag{50}$$

where σ_{in} and σ_{out}, respectively, denote the standard deviation of the entire original image and the processed image, and ϵ is a small constant used to prevent the denominator from being zero.

The signal-to-clutter ratio gain (SCRG) is a widely recognized metric employed to articulate the efficacy of the algorithm in augmenting the visibility of small targets. It is formally defined in the following manner:

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b + \varepsilon}, \quad SCRG = \frac{SCR_{\text{out}}}{SCR_{\text{in}}}. \tag{51}$$

where μ_t , μ_b, and σ_b represent the average value of the target region, the pixel average value and standard deviation of the surrounding local neighborhood region, respectively. SCR_{in} and SCR_{out} characterize the signal-to-clutter ratio of the original and processed images, respectively.

In Figure 7, it can be seen that a, b, and c represent the length and width of the target region, and the size of the width of the neighborhood region, respectively. In the experiments detailed within this paper, specific parameter values were set as follows: a and b were both assigned a value of 5, while c was determined to be 30.

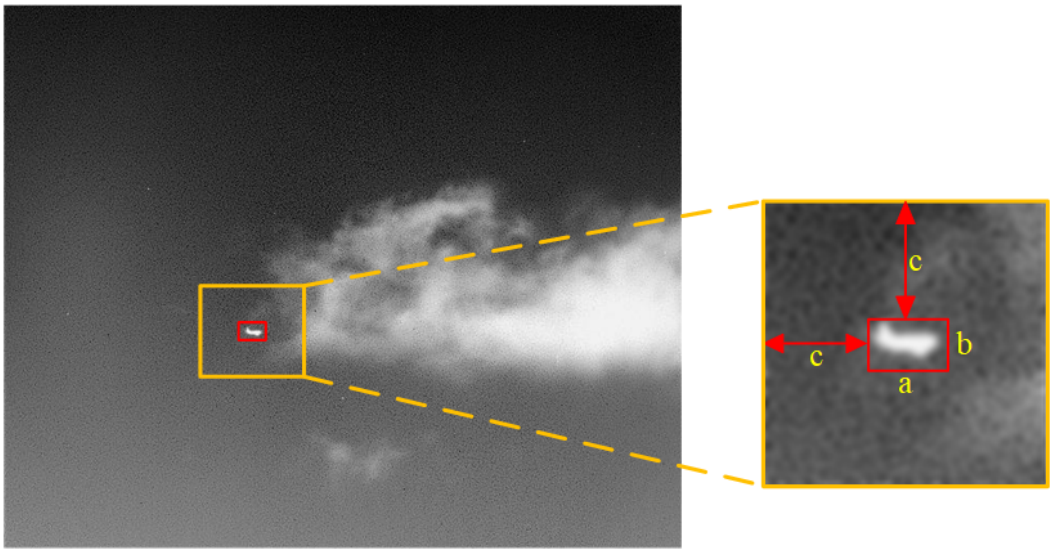


Figure 7. The neighboring background area of target.

The ROC curve is used in this study to comprehensively evaluate both the probability of detection and the false alarm rate, providing a thorough assessment of performance across various test scenarios. The definitions of the probability of detection and false alarm rate are as follows:

$$TPR = \frac{\text{number of detected true targets}}{\text{number of actual targets}} = \frac{TP}{TP + FN} \tag{52}$$

$$FPR = \frac{\text{number of detected false targets}}{\text{number of image pixels}} = \frac{FP}{AL} \tag{53}$$

TP, FN, FP, and AL represent the number of true targets detected, true targets not detected, false targets detected, and the total number of pixels in the analyzed image,

respectively. The ROC curve plots the False Positive Rate (FPR) on the x-axis against the True Positive Rate (TPR) on the y-axis, illustrating the algorithm's detection probability across various false alarm rates. The Area Under the Curve (AUC) derived from the ROC curve serves as a key performance metric, with a higher AUC value indicating better detection performance. A curve closer to the top left corner reflects the algorithm's superior ability to accurately identify true targets while minimizing false alarms.

4.2. Parameter Analysis

To investigate the robustness of the parameters and identify the optimal values for subsequent experiments, this section evaluates four critical parameters: patch size (ps), number of adjacent frames for tensor composition (p), penalty factor, μ, and balance coefficient, λ. Seq.1, Seq.2, and Seq.3 are used in a series of evaluation experiments.

4.2.1. Patch Size, ps

Patch size (ps) critically influences tensor construction. Larger ps values may compromise the low-rank assumption of the background by incorporating more image information. On the other hand, smaller ps values can lead to a spatial-temporal tensor with insufficient spatial neighborhood information, reducing the target's sparsity and, consequently, detection performance.

Building on previous research, this study conducted a comprehensive series of tests with patch sizes (ps) ranging from 10 to 40 in increments of 5. The results, presented in Figure 8(a1,b1,c1), demonstrate that the method adapts well to variations in ps, with optimal detection performance observed at a ps value of 25.

4.2.2. Adjacent Frames Number, p

The chosen value of p directly impacts both the algorithm's complexity and the amount of temporal information within the spatial-temporal tensor. This, in turn, determines the structure of the spatial-temporal patch tensor and the algorithm's complexity. Our empirical investigations tested p values of 2, 3, 4, 5, 6, 7, and 8. Analysis of the Receiver Operating Characteristic (ROC) curves shows that our method remains robust against variations in the number of adjacent frames, delivering strong detection performance in most scenarios. Based on a comprehensive analysis of the ROC curves, as shown in Figure 8(a2,b2,c2), a value of $p = 3$ was found to be optimal for achieving the best performance.

4.2.3. Penalty Factor, μ

The penalty factor, denoted as μ, regulates the balance between noise and sparse components by adjusting the shrinkage threshold. A lower penalty factor retains more background details, reducing the risk of non-target clutter being misclassified as sparse. However, this may also cause target details to merge with the background, leading to excessive shrinkage or even the loss of small targets. Conversely, a higher penalty factor preserves more target information but may allow more background clutter into the sparse tensor.

Based on insights from similar studies, our experiments tested values of 10, 20, 30, 40, and 50. The results, shown in Figure 8(a3,b3,c3), indicate that optimal performance occurs at $\mu = 20$.

4.2.4. Balance Factor, λ

The balance factor is also crucial. Taking λ_1 as an example, a higher λ_1 value helps recover more detailed textures in the background but may lead to more false alarms. Conversely, a lower λ_1 enhances the algorithm's sensitivity to weak targets, but this comes with an increased risk of false alarms.

In this paper, based on the experience from previous work [31,36,39], we set the balance factors as $\lambda_1 = 10^5 \lambda_L [\max(n_1, n_2) \times n_3]^{-1/2}$, where 10^5 is used to balance the parameter magnitude, and the inverse proportionality to the matrix dimensions is to accommodate

the tensor rank. Additionally, we set $\lambda_2 = \lambda_3 = \lambda_4 = 50\lambda_1$ to ensure that the designed optimization terms have sufficient effect. Thus, we only needed to adjust λ_L, significantly simplifying the parameter tuning process. Based on preliminary experiments, we tested λ_L values ranging from 0.5 to 2.0 with intervals of 0.3. The results in Figure 8 indicate that the best overall performance is achieved with $\lambda_L = 1.4$.

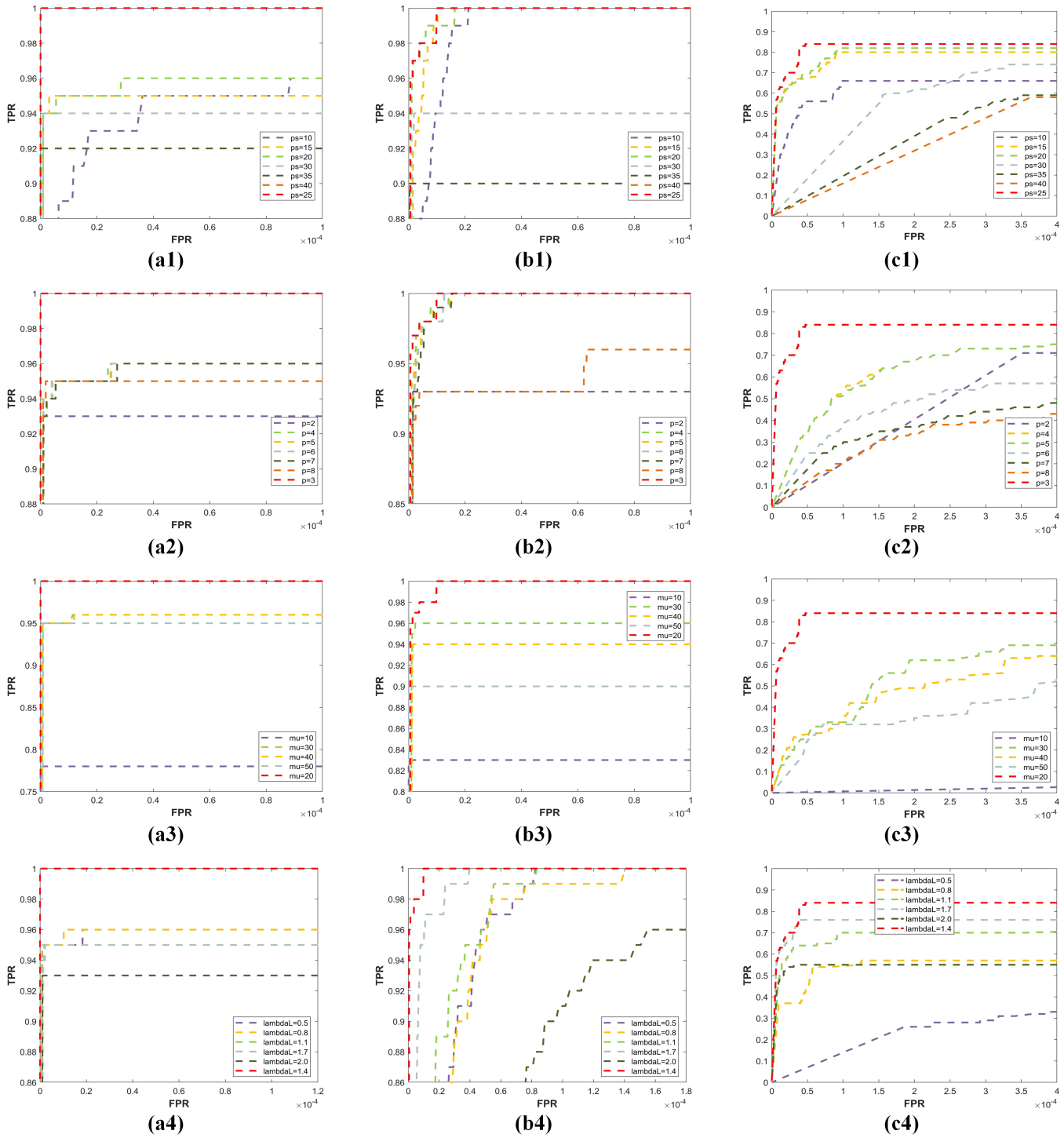


Figure 8. ROC curves for different parameters are displayed as follows: the first row for patch size results across three datasets, the second for adjacent frames, and the third and fourth for penalty and balance factors. Panels (**a1**–**a4**,**b1**–**b4**,**c1**–**c4**) correspond to each dataset's results, respectively.

4.3. Overall Performance Comparison

4.3.1. Qualitative Comparison

Figures 9 and 10 provide a qualitative comparison of detection results across six sequential datasets. Figure 9 shows grayscale images with true targets highlighted in red boxes, offering a clear visualization of the small target detection capabilities of different methods. The targets are magnified in the corners for clarity, and yellow boxes indicate missed detections by contrastive methods. Figure 10 displays 3D grayscale distribution maps of the detection results, visually demonstrating the background suppression capabilities of different methods. Identical annotations (e.g., (a1)) in both figures correspond to the same detection outcomes, simplifying references. Each figure includes only a single representative frame from each dataset for clarity.

TLLCM and DNGM both struggle in complex target environments, particularly in Seq.3–Seq.4, where point-like and linear clutter often lead to misidentifying clutter as targets. FKRW underperforms in complex backgrounds and with weak targets. As shown in (d3), (d4), and (d5), its detection efficacy and clutter suppression are visibly inferior to other methods.

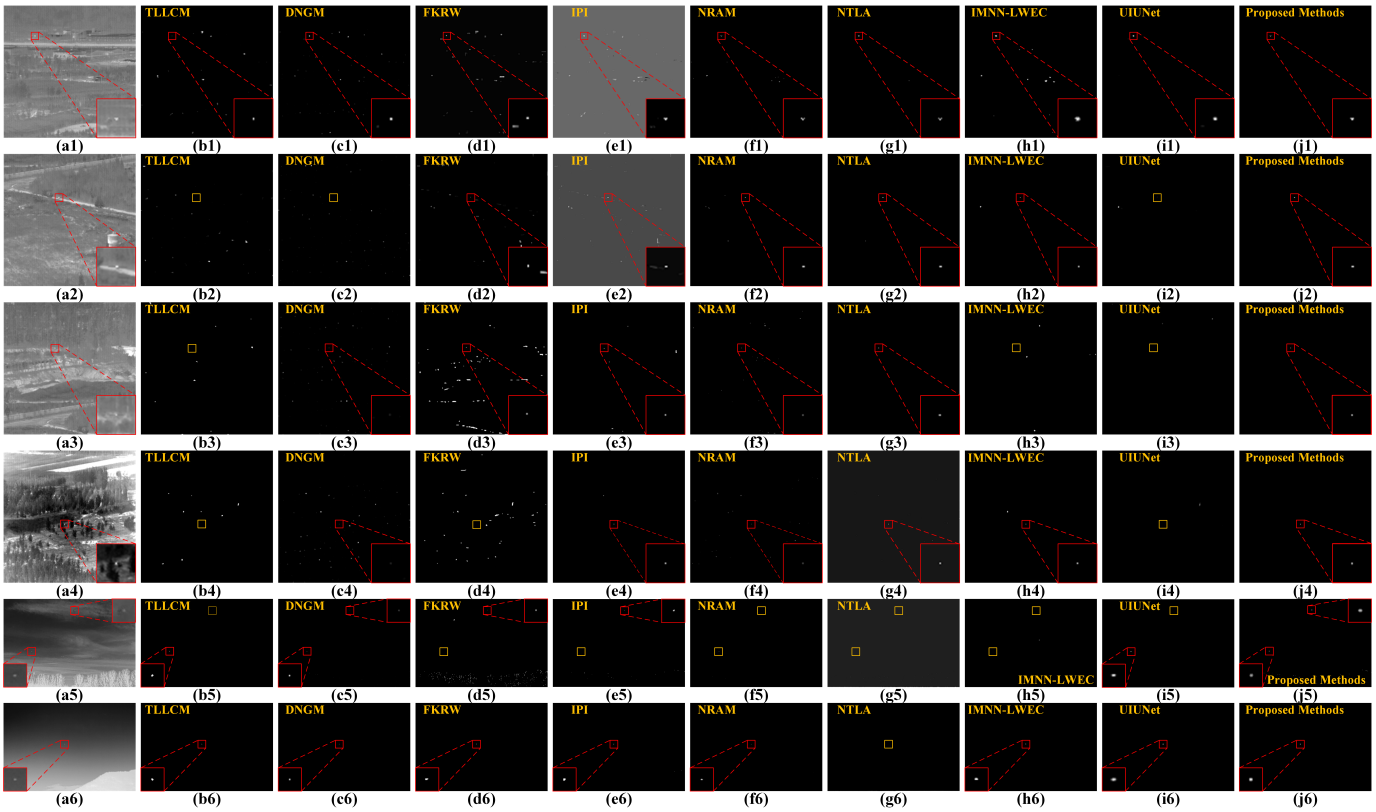


Figure 9. In analyzing Seq.1–Seq.6, targets are indicated with red boxes and magnified in corners, while yellow boxes mark undetected targets. Rows 1 to 6 correspond to the six sequences, and columns (**a1–a6**,**b1–b6**,**c1–c6**,**d1–d6**,**e1–e6**,**f1–f6**,**g1–g6**,**h1–h6**,**i1–i6**,**j1–j6**) denote different detection methods.

The IPI model and NRAM effectively detect targets amid clutter in Seq.1–Seq.4. However, the 3D-mesh and detection images show persistent flaws in clutter management, with background elements often misidentified as targets, as seen in (e3), (f4), and (g4). Additionally, these methods struggle with weaker targets in Seq.5 and Seq.6, as their nuclear norm-based background estimation fails to clearly distinguish faint targets from the background.

The multi-frame LRSD methods, NTLA and IMNN-LWEC, extend single-frame techniques by incorporating temporal information through tensor forms. Seq.1–Seq.4 show improved background suppression, as seen in the cleaner backgrounds in Figures 9 and 10.

However, like single-frame methods, they struggle with faint targets in Seq.5 and Seq.6, failing to detect them in (g5) and (h5). The deep learning-based UIUNet also shows strong background suppression capabilities. Nevertheless, as illustrated in (i2), (i3), and (i4), UIUNet experiences some degree of missed detection in the more complex scenarios represented by Seq.2–Seq.4.

In contrast, our proposed method shows greater robustness in complex scenes and better preserves the target object's shape. It effectively suppresses various types of clutter, making it highly effective for target detection in challenging scenarios. Compared with the other eight methods, our ILN-SSR method clearly excels in visual performance.

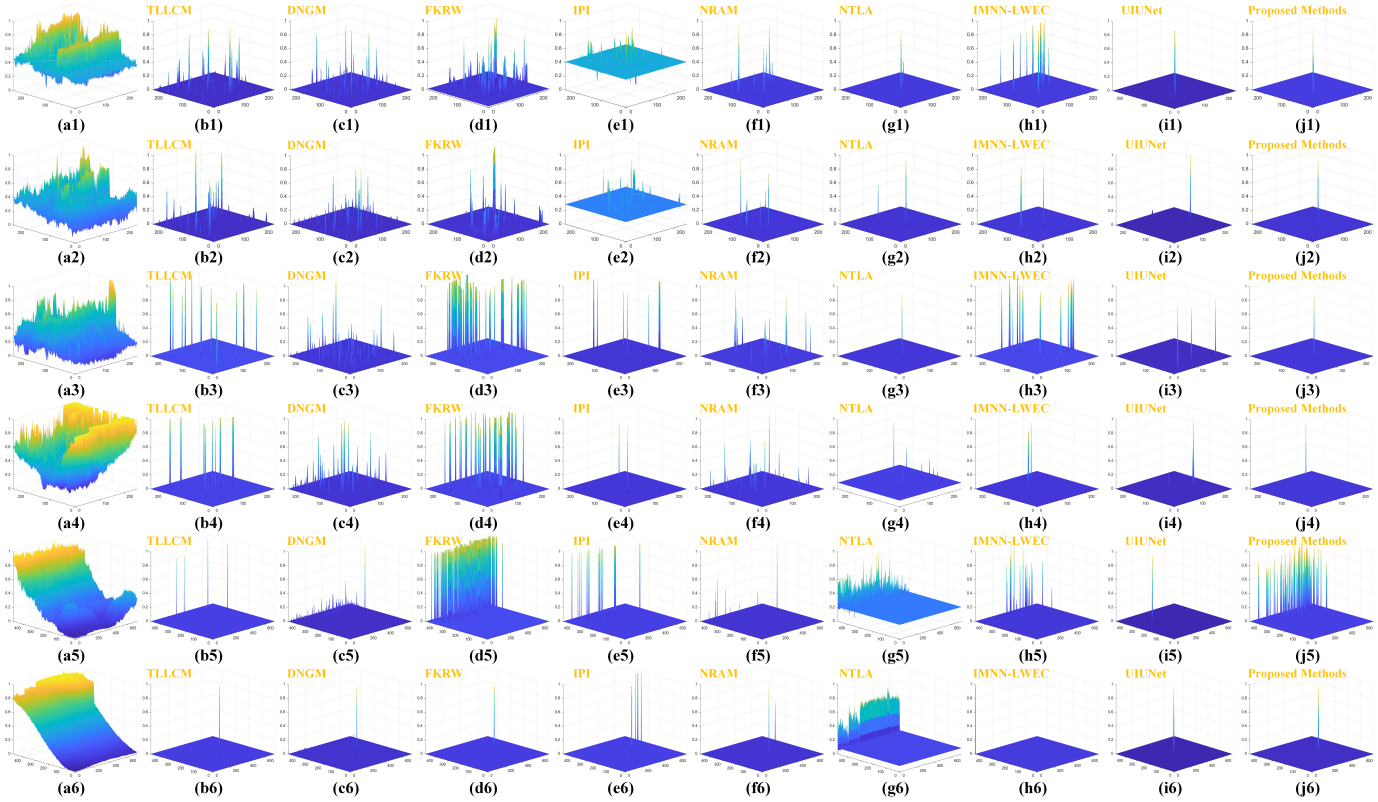


Figure 10. Comparing the intuitive 3D qualitative results for Seq.1–Seq.6, with image annotations corresponding to Figure 9.

4.3.2. Quantitative Comparison

This section employs AUC, SCRG, BSF, and ROC curves to conduct a comparative analysis of different methods, recognized as standard metrics in infrared small target detection. The findings are presented in Figure 11 and Table 3. In the table, the most outstanding results are highlighted in bold italics, while the second-best performances are underlined. It is pertinent to mention that the BSF and SCRG values are averages calculated over the entire sequence. In the case of SCRG, frames with missed detections are factored in, which could reduce this metric's value. Similarly, frames with a high frequency of false alarms are likely to influence the BSF.

TLLCM and DNGM exhibit strong detection capabilities in scenarios with clean target neighborhoods, as illustrated by the ROC curves in Seq.5 and Seq.6. However, their performance markedly declines with high background interference, as evidenced by lower ROC curves compared with those of most other methods in the study.

The detection results of FKRW, depicted by the green dashed lines in Figure 11a–f, are consistently absent from the upper half of the images, as discussed in the qualitative comparison section. This method consistently underperforms across multiple datasets.

The single-frame detection methods of IPI and NRAM are relatively effective in handling complex backgrounds, as demonstrated in Seq.3–Seq.4. However, they perform poorly when detecting faint targets in Seq.5–Seq.6.

Multi-frame detection methods like IMNN-LWEC and NLTA perform commendably in Seq.1 and Seq.2, where the target is relatively bright and the background is not overly complex. However, in Seq.3, these methods struggle to separate small targets from complex backgrounds, and in Seq.5 and Seq.6 their ability to detect dim and weak targets in simpler environments falls short of the best levels.

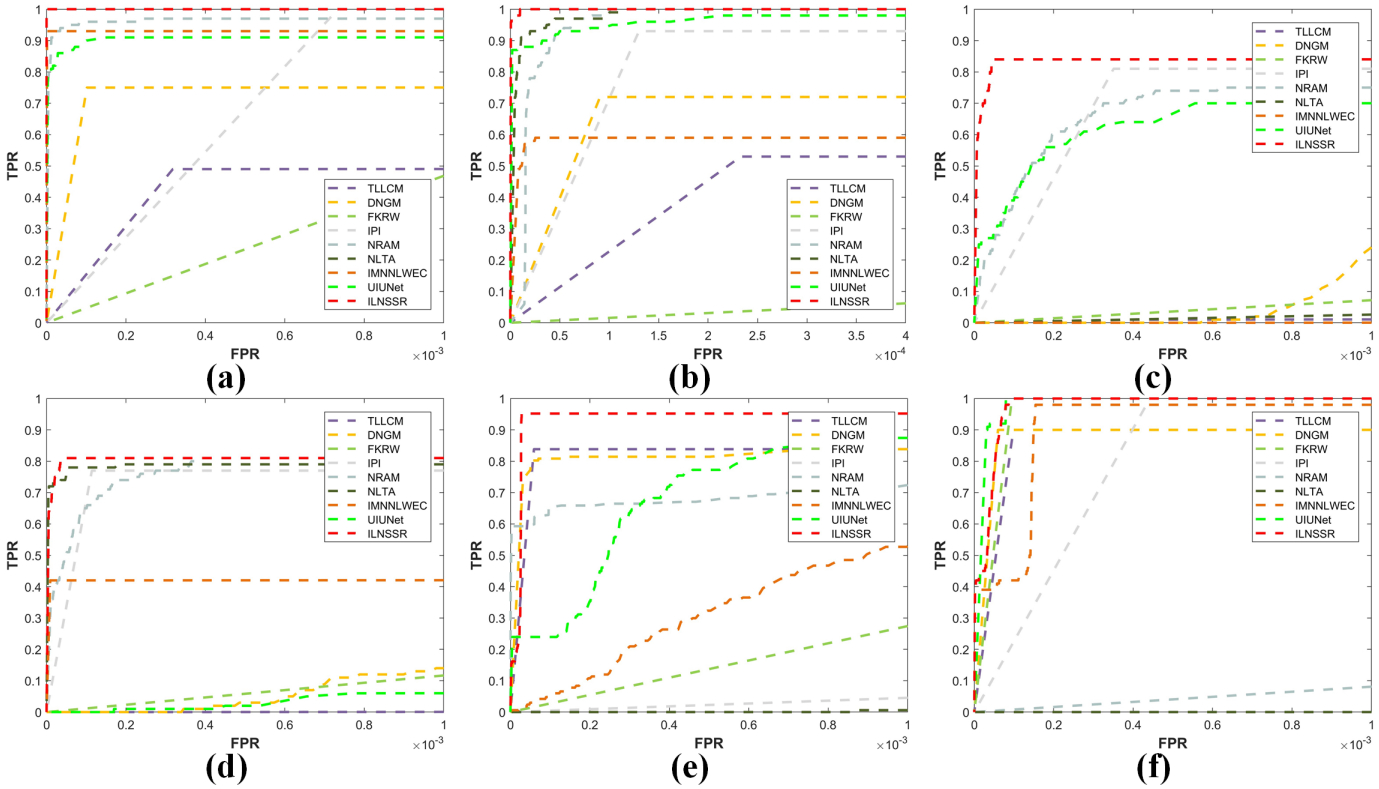


Figure 11. ROC curves achieved by different methods, (**a**–**f**) corresponding to Seq.1–Seq.6.

The deep learning-based UIUNet demonstrates good performance, achieving near-perfect detection in the less cluttered Seq.6. This aligns with the earlier qualitative comparison, showing that the method excels in background suppression. However, it falls short in the experiment with Seq.4, where the complex forest background makes it difficult for the model to learn the target's surrounding features, resulting in almost no detection of the target.

Analysis of Figure 11 shows that, among all evaluated methods, the ROC curve of our proposed method is nearest to the top left corner, achieving perfect detection in certain sequences. It is worth noting that, in Seq. 2 and Seq. 6, some single-frame methods yielded completely empty detection results for certain frames, leading to abnormally high BSF values that surpassed our method. However, this did not result in improved detection performance. In Seq. 4, the NRAM experimental results are close to those of ILN-SSR, with NRAM showing better SCRG metrics. However, while NRAM has fewer false alarms in the local neighborhood, it produces more global false alarms. Overall, considering the comprehensive data, our proposed method demonstrates the best detection performance. The above analysis underscores ILN-SSR's effectiveness in framework construction and the incorporation of various regularization components.

Table 3. AUC, SCRG and BSF of evaluated methods.

Sequence	Metrics	TLLCM	DNGM	FKRW	IPI	NRAM	NLTA	IMNNLWEC	UIUNet	Proposed
Seq.1	AUC	0.7448	0.8749	0.9789	0.9621	0.9849	0.9743	0.9649	0.9543	***1.0000***
	SCRG	1.47	19.13	1.18	24.26	22.08	***67.80***	42.80	53.63	***67.80***
	BSF	3.80	11.13	1.20	2.53	9.70	63.58	8.84	34.25	***69.09***
Seq.2	AUC	0.7648	0.8599	0.9023	0.9649	0.9997	0.9992	0.7949	0.9892	***0.9999***
	SCRG	46.38	16.51	0.63	87.37	32.44	69.00	35.63	75.72	***90.92***
	BSF	9.37	7.42	1.10	7.17	9.47	15.71	8.10	***58.30***	54.91
Seq.3	AUC	0.5044	0.8266	0.5981	0.8041	0.8241	0.7896	0.4992	0.8192	***0.8293***
	SCRG	0.04	6.09	2.24	23.21	14.89	47.23	0.00	43.02	***53.77***
	BSF	4.14	6.54	1.58	4.61	7.89	31.21	4.66	23.68	***56.02***
Seq.4	AUC	0.4994	0.7799	0.6284	0.8842	0.9042	0.8739	0.7094	0.5292	***0.9043***
	SCRG	0.00	2.90	0.37	16.05	***85.90***	50.10	4.67	14.28	81.96
	BSF	11.60	14.91	4.87	21.42	23.53	52.06	25.30	33.40	***71.46***
Seq.5	AUC	0.9183	0.9188	0.9323	0.9536	0.9214	0.6378	0.8434	0.9366	***0.9757***
	SCRG	48.50	46.97	45.12	59.48	51.94	70.46	52.38	59.18	***70.62***
	BSF	24.21	25.51	3.90	12.25	19.43	9.71	15.56	43.70	***61.38***
Seq.6	AUC	0.9498	0.9498	0.9996	0.9992	0.9848	0.6558	0.9898	***0.9998***	***0.9998***
	SCRG	41.10	36.31	39.69	46.51	48.11	31.26	48.75	28.21	***49.23***
	BSF	96.32	99.72	105.23	33.37	***121.67***	7.85	113.78	86.20	101.66

4.3.3. Ablation Study

To ascertain the efficacy of each module within ILN-SSR, an ablation study was conducted using dataset 2, with the outcomes detailed in the Table 4. Notably, the best-performing segments in the results are highlighted in bold italics. The baseline of the experiment is established by the LRSD framework, with nuclear norm and no sparse refinement, which constructs a spatial-temporal tensor using three consecutive frames from Section 3.1.

In this baseline scenario, ILN, the Weight Prior Map, and the linear sparse term were gradually added or replaced to comprehensively evaluate the performance of each module within ILN-SSR.

Table 4. Ablation study results (wherein ✓ indicates that this setting is included in the corresponding ablation experiment options).

LRSD Framework	ILN Surrogate	Prior Weight Map	Sparse Structure Refinement		AUC	SCRG	BSF
			Linear Sparse Term	Target-Traits Term			
✓					0.9662	23.12	3.31
✓	✓				0.9990	31.25	3.37
✓		✓			0.9994	32.21	6.21
✓	✓	✓			0.9998	32.36	16.77
✓	✓	✓	✓		***0.9999***	41.65	19.72
✓	✓	✓	✓	✓	***0.9999***	***90.92***	***54.91***

Bold-italic values indicate the best performance in each column. Underlined values represent the second-best performance in each column.

4.3.4. Analysis of Running Time

In addition to evaluation metrics such as ROC curves, BSF, and SCRG, runtime plays a vital role in performance assessment. Table 4 presents the average runtime per frame for nine detection algorithms across six sequences. Typically, single-frame detection methods outperform multi-frame methods in processing speed, as the latter incorporate temporal information between frames. LRSD-based detection methods often strike a balance between improved performance and increased computational complexity due to intricate optimization calculations. Notably, among the evaluated methods, IMNN-LWEC exhibits the longest processing time in Seq.1–Seq.4, while the IPI method takes the longest in Seq.5–Seq.6.

Considering the additional steps involved in our algorithm, such as extracting sparse components and more precise estimation of the restored background tensor, the increment in runtime is deemed justifiable. When correlating the findings from Tables 4 and 5, it is evident that our method surpasses other competing approaches in overall efficacy. Looking forward, there is potential to further optimize the algorithm or leverage GPUs to improve the real-time performance capabilities of our approach.

Table 5. Running time analysis (running time of different methods).

Methods	Seq.1	Seq.2	Seq.3	Seq.4	Seq.5	Seq.6
TLLCM	25.90 s	25.10 s	25.13 s	25.46 s	112.35 s	119.24 s
DNGM	3.85 s	3.83 s	3.88 s	3.80 s	7.75 s	7.67 s
FKRW	8.02 s	7.92 s	7.99 s	7.88 s	21.33 s	21.29 s
IPI	163.60 s	159.27 s	163.79 s	166.56 s	1638.48 s	1533.58 s
NRAM	74.50 s	81.43 s	80.18 s	68.56 s	760.21 s	756.07 s
NLTA	111.27 s	94.32 s	89.74 s	91.14 s	804.64 s	809.52 s
IMNN-LWEC	206.54 s	165.43 s	167.24 s	163.13 s	830.71 s	826.59 s
Proposed	160.33 s	150.46 s	149.79 s	151.23 s	707.16 s	715.26 s

5. Discussion

By adopting a framework that integrates an improved logarithmic norm with a sparse structure, the proposed method thoroughly considers various causes of false alarms in infrared images, demonstrating excellent performance not only in target enhancement but also in background suppression.

The experimental section of this paper conducted a series of experiments, including robustness to various scenarios, target enhancement, background suppression, detection ability, and computation time, to compare the proposed method with other benchmark methods. The experimental results indicate that the proposed method outperforms seven representative state-of-the-art methods, including TLLCM, DNGM, IPI, FKRW, NRAM, NLTA, and IMNNLWEC.

Optimized methods remain an important direction for solving tasks such as target detection. The authors of this paper believe that, in addition to describing images through different regularization terms, future research directions will involve how to decompose images into tensor representations and how to introduce the concept of parallel computing into optimization.

6. Conclusions

In this article, we introduce a spatial-temporal framework named improved logarithmic norm joint sparse structure refinement for infrared small target detection. This algorithm is designed to accurately represent tensor background rank and robustly extract sparse components. For tensor rank representation, we employ an improved logarithmic

norm to more precisely estimate the nonconvex tensor rank. Regarding the extraction of sparse components, regularization factors are utilized, which are attuned to target characteristics and linear noise. Moreover, a local structure tensor is incorporated as prior information for pre-detection optimization.

The target detection challenge is addressed by optimizing the ILN-SSR framework, employing a layered analytical approach based on the alternating direction method of multipliers. This method facilitates a step-by-step optimization of each factor, thereby resolving the optimization problem effectively. Comprehensive experimental evaluations demonstrate that our algorithm excels in detecting moving targets amidst complex backgrounds, proficiently suppressing and differentiating background clutter. The mainstream evaluation metrics underscore our algorithm's superior and consistent performance relative to other algorithms.

Nevertheless, the proposed framework has its limitations. Primarily, it is time-intensive, which highlights the need for more efficient optimization strategies or the integration of parallel computing techniques to enhance its performance. Additionally, while our approach for constructing spatial-temporal tensors has demonstrated effectiveness in extracting temporal information throughout the experiments, it reveals shortcomings in establishing robust inter-frame correlations. This limitation points to an area that requires deeper exploration, as strengthening inter-frame correlations could lead to more accurate and reliable target detection. Future research could also benefit from investigating advanced tensor decomposition techniques and exploring ways to better balance the trade-off between computational cost and detection accuracy, ensuring the framework can be applied to real-time or large-scale scenarios. Moreover, integrating the framework with neural network-based methods, particularly those employing deep unfolding techniques, might provide further improvements in both efficiency and detection precision.

Author Contributions: Conceptualization, F.C. and L.L. (Liyuan Li); methodology L.L. (Liqi Liu) and X.N.; software, L.L. (Liqi Liu) and R.Z.; validation, L.L. (Liqi Liu) and J.M.; formal analysis, L.L. (Liqi Liu) and X.S.; investigation, L.L. (Liqi Liu) and X.N.; resources, L.L. (Liqi Liu) and X.S.; data curation, L.L. (Liqi Liu); writing—original draft preparation, L.L. (Liqi Liu); writing—review and editing, L.L. (Liyuan Li); supervision, F.C.; project administration, F.C. and L.L. (Liyuan Li); funding acquisition F.C. and L.L. (Liyuan Li). All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by National Natural Science Foundation of China (Grant No. 61975222) and CASEarth Minisatellite Thermal Infrared Spectrometer Project (XDA19010102).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Abbreviations	Full Names
ADMM	Alternating direction method of multipliers
AUC	Area Under the Curve
BSF	Background Suppression Factor
FPR	False Positive Rate
ILN-SSR	Improved logarithmic norm and sparse structure refinement
IRSTD	Infrared small target detection
LRSD	Low-rank and sparse decomposition
nFFT	Non-uniform fast Fourier transform
ROC	Receiver Operating Characteristic
SCR	Signal-to-clutter ratio
SOTA	State of the art
SVD	Singular value decomposition
T-SVD	Tensor singular value decomposition

TPR	True Positive Rate
$\mathcal{B}$	Tensor composed of background patches
$\mathcal{E}_1$	Tensor representing column sparsity
$\mathcal{E}_2$	Tensor representing row sparsity
$\mathcal{N}$	Tensor composed of noise patches
$\mathcal{P}$	Sparse structures related to non-target attributes
$\mathcal{T}$	Tensor composed of target patches
$\mathcal{V}$	Auxiliary components used for optimization
$\mathcal{W}$	Weight tensors
$\mathcal{Z}$	Auxiliary components used for optimization

References

1. Paugam, R.; Wooster, M.J.; Roberts, G. Use of handheld thermal imager data for airborne mapping of fire radiative power and energy and flame front rate of spread. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 3385–3399. [CrossRef]
2. Price, J.; Maraviglia, C.; Seisler, W.; Williams, E.; Pauli, M. System capabilities, requirements and design of the GDL gunfire detection and location system. In Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop (AIPR'04), Washington, DC, USA, 13–15 October 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 257–262.
3. Zhang, X.; Wu, N.; Liu, C.; Tong, L. Ballistic missile launch position estimation based on the first point measurement data of space infrared sensor. In Proceedings of the 2022 4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP), Hangzhou, China, 8–10 July 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 271–274.
4. Li, L.; Zhou, X.; Hu, Z.; Gao, L.; Li, X.; Ni, X.; Chen, F. On-orbit monitoring flying aircraft day and night based on SDGSAT-1 thermal infrared dataset. *Remote Sens. Environ.* **2023**, *298*, 113840. [CrossRef]
5. Zhang, W.; Cong, M.; Wang, L. Algorithms for optical weak small targets detection and tracking. In Proceedings of the International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003, Nanjing, China, 14–17 December 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 1, pp. 643–647.
6. Tartakovsky, A.G.; Kligys, S.; Petrov, A. Adaptive sequential algorithms for detecting targets in a heavy IR clutter. In Proceedings of the Signal and Data Processing of Small Targets 1999, Orlando, FL, USA, 20–22 July 1999; SPIE: Bellingham, WA, USA, 1999; Volume 3809, pp. 119–130.
7. 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. Image Process.* **1993**, *2*, 327–340. [CrossRef] [PubMed]
8. Deshpande, S.D.; Er, M.H.; Venkateswarlu, R.; Chan, P. Max-mean and max-median filters for detection of small targets. In Proceedings of the Signal and Data Processing of Small Targets 1999, Orlando, FL, USA, 20–22 July 1999; SPIE: Bellingham, WA, USA, 1999; Volume 3809, pp. 74–83.
9. Farajzadeh, M.; Mahmoodi, A.; Arvan, M.R. Detection of small target based on morphological filters. In Proceedings of the 20th Iranian Conference on Electrical Engineering (ICEE2012), Tehran, Iran, 15–17 May 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 1097–1101.
10. Qin, Y.; Bruzzone, L.; Gao, C.; Li, B. Infrared small target detection based on facet kernel and random walker. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7104–7118. [CrossRef]
11. Lu, L. Research on infrared small target detection and tracking algorithms based on wavelet transformation. *Sens. Transducers* **2013**, *156*, 116.
12. Wang, H.; Xin, Y. Wavelet-based contourlet transform and kurtosis map for infrared small target detection in complex background. *Sensors* **2020**, *20*, 755. [CrossRef]
13. Qi, S.; Ma, J.; Li, H.; Zhang, S.; Tian, J. Infrared small target enhancement via phase spectrum of quaternion Fourier transform. *Infrared Phys. Technol.* **2014**, *62*, 50–58. [CrossRef]
14. Kim, S.; Yang, Y.; Lee, J.; Park, Y. Small target detection utilizing robust methods of the human visual system for IRST. *J. Infrared Millim. Terahertz Waves* **2009**, *30*, 994–1011. [CrossRef]
15. Chen, C.P.; Li, H.; Wei, Y.; Xia, T.; Tang, Y.Y. A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 574–581. [CrossRef]
16. Qin, Y.; Li, B. Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1890–1894. [CrossRef]
17. Han, J.; Liang, K.; Zhou, B.; Zhu, X.; Zhao, J.; Zhao, L. Infrared small target detection utilizing the multiscale relative local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 612–616.
18. Wei, Y.; You, X.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recogn.* **2016**, *58*, 216–226.
19. Han, J.; Moradi, S.; Faramarzi, I.; Liu, C.; Zhang, H.; Zhao, Q. A local contrast method for infrared small-target detection utilizing a tri-layer window. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1822–1826.
20. Du, P.; Hamdulla, A. Infrared small target detection using homogeneity-weighted local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 514–518.

21. Wu, L.; Ma, Y.; Fan, F.; Wu, M.; Huang, J. A double-neighborhood gradient method for infrared small target detection. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1476–1480.
22. Wang, H.; Zhou, L.; Wang, L. Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8509–8518.
23. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Attentional local contrast networks for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9813–9824.
24. Dai, Y.; Li, X.; Zhou, F.; Qian, Y.; Chen, Y.; Yang, J. One-stage cascade refinement networks for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5000917.
25. Wu, X.; Hong, D.; Chanussot, J. UIU-Net: U-Net in U-Net for infrared small object detection. *IEEE Trans. Image Process.* **2022**, *32*, 364–376.
26. Li, Y.; Xu, Q.; He, Z.; Li, W. Progressive task-based universal network for raw infrared remote sensing imagery ship detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5610013.
27. Li, Y.; Xu, Q.; Kong, Z.; Li, W. MULS-Net: A Multilevel Supervised Network for Ship Tracking From Low-Resolution Remote-Sensing Image Sequences. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5624214.
28. Tong, X.; Su, S.; Wu, P.; Guo, R.; Wei, J.; Zuo, Z.; Sun, B. MSAFFNet: A multiscale label-supervised attention feature fusion network for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5002616. [CrossRef]
29. Zhang, M.; Yue, K.; Li, B.; Guo, J.; Li, Y.; Gao, X. Single-frame infrared small target detection via gaussian curvature inspired network. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5005013. [CrossRef]
30. Kou, R.; Wang, C.; Peng, Z.; Zhao, Z.; Chen, Y.; Han, J.; Huang, F.; Yu, Y.; Fu, Q. Infrared small target segmentation networks: A survey. *Pattern Recogn.* **2023**, *143*, 109788. [CrossRef]
31. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef]
32. Zhang, L.; Peng, Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sens.* **2019**, *11*, 382. [CrossRef]
33. Wang, G.; Tao, B.; Kong, X.; Peng, Z. Infrared small target detection using nonoverlapping patch spatial–temporal tensor factorization with capped nuclear norm regularization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–17. [CrossRef]
34. Guan, X.; Zhang, L.; Huang, S.; Peng, Z. Infrared small target detection via non-convex tensor rank surrogate joint local contrast energy. *Remote Sens.* **2020**, *12*, 1520. [CrossRef]
35. Zhu, H.; Ni, H.; Liu, S.; Xu, G.; Deng, L. Tnlrs: Target-aware non-local low-rank modeling with saliency filtering regularization for infrared small target detection. *IEEE Trans. Image Process.* **2020**, *29*, 9546–9558. [CrossRef]
36. Liu, T.; Yang, J.; Li, B.; Xiao, C.; Sun, Y.; Wang, Y.; An, W. Nonconvex tensor low-rank approximation for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5614718. [CrossRef]
37. Zhang, P.; Zhang, L.; Wang, X.; Shen, F.; Pu, T.; Fei, C. Edge and corner awareness-based spatial–temporal tensor model for infrared small-target detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 10708–10724. [CrossRef]
38. Zhang, Z.; Ely, G.; Aeron, S.; Hao, N.; Kilmer, M. Novel methods for multilinear data completion and de-noising based on tensor-SVD. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3842–3849.
39. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm. *Remote Sens.* **2018**, *10*, 1821. [CrossRef]
40. Luo, Y.; Li, X.; Chen, S.; Xia, C.; Zhao, L. IMNN-LWEC: A novel infrared small target detection based on spatial–temporal tensor model. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5004022. [CrossRef]
41. Kong, X.; Yang, C.; Cao, S.; Li, C.; Peng, Z. Infrared small target detection via nonconvex tensor fibered rank approximation. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5000321. [CrossRef]
42. Chen, L.; Jiang, X.; Liu, X.; Zhou, Z. Logarithmic norm regularized low-rank factorization for matrix and tensor completion. *IEEE Trans. Image Process.* **2021**, *30*, 3434–3449. [CrossRef]
43. Li, M.; Li, X.; Jiang, Y.; Zhang, J.; Luo, H.; Yin, S. Explainable multi-instance and multi-task learning for COVID-19 diagnosis and lesion segmentation in CT images. *Knowl. Based Syst.* **2022**, *252*, 109278. [CrossRef]
44. Liu, J.; Wu, J.; Xu, M.; Huang, Y. Plug-and-play-based algorithm for mixed noise removal with the logarithm norm approximation model. *Mathematics* **2022**, *10*, 3810. [CrossRef]
45. Kilmer, M.E.; Braman, K.; Hao, N.; Hoover, R.C. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 148–172. [CrossRef]
46. Kilmer, M.E.; Martin, C.D. Factorization strategies for third-order tensors. *Linear Algebra Its Appl.* **2011**, *435*, 641–658. [CrossRef]
47. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 925–938. [CrossRef]
48. Golub, G.H.; Van Loan, C.F. *Matrix Computations*; JHU Press: Baltimore, MD, USA, 2013.
49. Gu, Y.; Wang, C.; Liu, B.; Zhang, Y. A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 469–473. [CrossRef]

50. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 5249–5257.
51. Kang, Z.; Peng, C.; Cheng, Q. Top-n recommender system via matrix completion. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30.
52. Zhou, P.; Lu, C.; Lin, Z.; Zhang, C. Tensor factorization for low-rank tensor completion. *IEEE Trans. Image Process.* **2017**, *27*, 1152–1163. [CrossRef] [PubMed]
53. Gao, C.Q.; Tian, J.; Wang, P. Generalised-structure-tensor-based infrared small target detection. *Electron. Lett.* **2008**, *44*, 1. [CrossRef]
54. Candes, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted ℓ 1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [CrossRef]
55. Daubechies, I.; Defrise, M.; De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.* **2004**, *57*, 1413–1457. [CrossRef]
56. Helou, M.E.; Dümbgen, F.; Achanta, R.; Süsstrunk, S. Fourier-domain optimization for image processing. *arXiv* **2018**, arXiv:1809.04187.
57. Duijndam, A.; Schonewille, M. Nonuniform fast Fourier transform. *Geophysics* **1999**, *64*, 539–551. [CrossRef]
58. Hui, B.; Song, Z.; Fan, H.; Zhong, P.; Hu, W.; Zhang, X.; Lin, J.; Su, H.; Jin, W.; Zhang, Y.; et al. A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background. *Sci. Data Bank* **2019**, *5*, 4. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.