



Article

An Enhanced Image Patch Tensor Decomposition for Infrared Small Target Detection

Ziling Lu ¹, Zhenghua Huang ², Qiong Song ^{1,*}, Kun Bai ³ and Zhengtao Li ⁴¹ School of Computer Science, Northeast Electric Power University, Jilin 132012, China² Artificial Intelligence School, Wuchang University of Technology, Wuhan 430223, China³ Xi'an Modern Control Technology Research Institute, Xi'an 710065, China⁴ College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300382, China

* Correspondence: 20192918@neepu.edu.cn

Abstract: Infrared small-target detection is a key technology for the infrared search and track system (IRST), but some problems still exist, such as false detections in complex backgrounds and clutter. To solve these problems, a novel image patch tensor (IPT) model for infrared small-target detection is proposed. First, to better estimate the background component, we utilize the Laplace operator to approximate the background tensor rank. Secondly, we combined local gradient features and highlighted area indicators to model the local targets prior, which can effectively suppress the complex background clutter. The proposed model was solved by the alternating direction method of multipliers (ADMM). The experimental results on various scenes show that our model achieves an excellent performance in suppressing strong edge clutter and estimating small targets.

Keywords: highlighted area indicator; small target detection; low rank sparse decomposition; infrared image; gradient feature



Citation: Lu, Z.; Huang, Z.; Song, Q.; Bai, K.; Li, Z. An Enhanced Image Patch Tensor Decomposition for Infrared Small Target Detection. *Remote Sens.* **2022**, *14*, 6044. <https://doi.org/10.3390/rs14236044>

Academic Editor: Salah Bourennane

Received: 21 October 2022

Accepted: 25 November 2022

Published: 29 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Infrared (IR) small-target detection plays a crucial role in IRST and is a key technology for different military and civilian applications, such as air raid warnings, maritime rescues, and power equipment fault detection. However, the detection performance in practical applications is not always satisfactory for the following reasons. (1) The small target at a long distance usually has less texture and shape features that can be used for detection. (2) The targets have low intensity under the interference of complex backgrounds, such as clouds and solar radiation. Therefore, the way to effectively suppress background and improve the target detection performance for the IRST is still a challenging task.

In general, track-before-detect (TBD) and detect-before-track (DBT) are two main methods for small-target detection. TBD uses multiframe to detect targets, and typical methodologies include 3D matched filtering [1], dynamic programming [2], and pipeline filtering [3]. In real scenes, the TBD methods usually require huge amounts of computational power to process the sequence to detect the target, which makes them less time-efficient. In contrast, DBT uses prior information of the targets and background to detect a target in a single frame, so the algorithm's computation is faster. DBT algorithms can be broadly classified into three types.

(1) Background suppression-based methods

The background suppression-based methods consider the small target to be an independent and distinct component from the background. They usually use some filters to suppress the background and then utilize a threshold function to extract small targets. Early on, researchers used bilateral filtering [4], maximum mean filtering, and maximum median filtering [5] to filter an original IR image to estimate the background, and then they used the obtained target image to segment the target. The minimum mean squared difference

filter [6] estimates a target image by the weighted average of the neighbouring pixels. In addition, some morphological theory methods, such as Top-hat, have often been used to separate small targets from the background [7–9]. However, most of these methods require prior knowledge of small targets and may be affected by noise or edge clutter. Recently, some researchers have improved some filters [10,11] to estimate the background components for detecting the small target. However, this type of method could not suppress complex backgrounds well.

(2) Human visual system (HVS)-based methods

HVS-based methods consider a small target to be visually more prominent than the surrounding background. Taking advantage of this property, Chen et al. [12] proposed the local contrast measure (LCM) to measure the difference between the center patch and surrounding areas. Subsequently, a series of improved LCMs were proposed. Qin et al. [13] enhanced the target by using a new LCM so that the target is not influenced by the highlighted background (NLCM). The authors in [14] combined the local intensity and gradient (LIG) to enhance the target's contrast and eliminate background clutter. Some methods, such as DNGM [15], TTLCM [16], and VARD [17] use multilayer window models to enhance small targets. To address the interference of different background clutters, Bai et al. [18] combined derivative entropy and LCM (DECM) to enhance the target while suppressing noise. Generally, HVS-based methods can discriminate the target well when the background is relatively smooth. However, the HVS-based method could not suppress some strong edges well.

(3) Optimization model-based methods

This type of method usually formulates a model based on the sparsity of small targets in an image and the nonlocal correlation of the background. Gao et al. [19] first designed an infrared image patch model (IPI), which assumes that the background component is low-rank and the target is sparse. They formulated an optimization model based on sparse and low-rank matrix recovery to separate small targets. However, the limitations of nuclear norm minimization (NNM) in the IPI model can lead to the target overshrinkage problem [20]. To solve the overshrinking problem, Dai et al. proposed weighting each column of the target image patch to obtain a global weight and named this method the WIPI [21] model. However, WIPI is slightly computationally intensive. Zhang et al. employed the $L_{2,1}$ norm to suppress sparse strong edges in the nonconvex rank approximation minimization (NRAM) model [22]. In [23], the authors utilized the L_p norm to constrain sparse targets and proposed nonconvex optimization with the L_p norm constraint (NOLC) model. By adding total variation (TV) regularization, Wang et al. [24] combined total variation regularization and principal component pursuit (TV-PCP) to effectively suppress edge clutter in an image. Several multisubspace learning models have been designed, such as the stable multisubspace learning (SMSL) method [25] and self-regularized weighted sparse (SRWS) model [26]. Dai et al. [27] decomposed the input in the tensor domain rather than the matrix domain and proposed a reweighted infrared patch tensor (RIPT) model. The RIPT model can achieve relatively good performance because the tensor model can make better use of the interpixel resultant information. Zhang et al. [28] utilized the partial sum of the tensor nuclear norm (PSTNN) to constrain the low-rank background tensor, and this technique can detect the small target more effectively. Guan et al. [29] integrated local contrast energy into the optimization process to eliminate structured edges.

With the development of deep learning, many researchers formulate the neural network for target detection and location [30–34], such as RCNN [35], Faster-RCNN [36], YOLO [37], SSD [38], and FPN [39]. In recent years, some network frameworks are also proposed for the small-target detection task. In [40], Wang et al. proposed a coarse-to-fine network for small-target detection. They first applied a region potential network to estimate coarse target regions, then they used the transformer encoder to model the interior relation for pixels in the coarse regions, and finally they predicted the target by the attention-aware features. In order to preserve the target feature in the deep layers, Li et al. [41] proposed a

dense nested attention network (DNANet). Shi et al. [42] proposed an end-to-end detection network based on the denoising autoencoder network and convolutional neural network in which they treat small-target detection problems as noise-removal problems. Although the deep learning-based small-target detection methods could achieve some excellent results, the huge dataset that cover various scenes needed by this technique are difficult to be obtained.

In general, the background components in an image are supposed to be slowly transitioning, which indicates that those adjacent image patches are greatly correlated. Thus, the background in an image can extract a powerful low-rank property [43]. With regard to the IPT-based model, accurately approximating the background rank is one main issue [44]. The RIPT model uses the sum of the nuclear norm (SNN) to approximately estimate the rank of the background tensor, whereas the unfolding implementation destroys the patch structure, so the SNN cannot accurately approximate the tensor rank. The tensor nuclear norm (TNN) is another common technique that approximates the background tensor's rank in research and applications [45,46]. However, all singular values in the original TNN are equally important, and in fact different singular values in an image have different importance and physical significance, so different treatment should be given to different singular values. In PSTNN [28], the authors truncate some small singular values, and the large singular values remain with the same weight. To more accurately approximate the background rank, we propose a Laplace function-based approximation method, which could intelligently designate different weights to unequal singular values.

The local and nonlocal priors are both beneficial and complementary for the infrared small-target detection mission [27]. Those methods that focus on local features are not sufficient to distinguish between background and targets [47–50]. The low-rank sparse matrix recovery model uses nonlocal features, but it is very sensitive to those edges that exist in the background and has poor suppression ability. The strong edges with sparsity cannot be well retained in the background by the low rank constraint. In fact, strong edges have distinct local features and can be suppressed by local prior information. Thus, a new IPT model based on local and nonlocal priors is proposed in this paper.

To better detect small targets in complex scenes, in this paper, we present a new IPT model based on gradient features and edge and highlighted area indicators. Our paper makes the following contributions.

- (1) To estimate the background more accurately, we use the Laplace operator to approximate the background tensor rank; it assigns different weights to different singular values.
- (2) A new small target aware prior. We combine local gradient features and edge and highlighted area indicators to enhance the small targets and suppress those sharp edges.
- (3) We introduce the Laplace operator and structural prior into the IPT model and use the ADMM to solve our model.

The arrangement of the rest of the content is as follows. In Section 2, we give some necessary mathematical notations and definitions. Section 3 presents the proposed model, including the construction of the local prior, the introduction of the Laplace approximation, and a detailed description of the optimization process. Section 4 describes the comparison experiments and provides a qualitative and quantitative evaluation. Finally, the conclusions are given in Section 5.

2. Notations

The symbols and the corresponding explanations that appear in this paper are listed in Table 1. Knowledge about tensor, definitions, and theorems can be found in the Appendix A.

Table 1. Mathematical symbols.

Notation	Instruction
$a/a/A/A$	scalar/vector/matrix/tensor
$\mathcal{A}(i,:)/\mathcal{A}(:,i)/\mathcal{A}(:,:,i)$ or $\mathcal{A}^{(i)}$	the i th horizontal slice / lateral slice / frontal slice of tensor \mathcal{A}
\mathcal{A}^i	the i -th iteration of \mathcal{A}
$\ \mathcal{A}\ _0$	l_0 norm of tensor \mathcal{A} , which is the number of nonzero elements
$\ \mathcal{A}\ _1$	l_1 norm of tensor \mathcal{A} , which is the absolute sum of all elements in \mathcal{A}
$\ \mathcal{A}\ _F$	Frobenius norm of tensor \mathcal{A} , which is the root of the sum of the absolute values of the squares of the tensor elements
$\ \mathcal{A}\ _*$	nuclear norm of tensor \mathcal{A} , which is the sum of all the singular values
$\overline{\mathcal{A}}=\text{fft}(\mathcal{A}, [], 3)/\mathcal{A}=\text{ifft}(\overline{\mathcal{A}}, [], 3)$	fast Fourier transform of \mathcal{A} / inverse Fourier transform of $\overline{\mathcal{A}}$

3. Proposed Model

3.1. Image Patch Tensor (IPT) Model

In general, an infrared image that contained a small target could be decomposed into three components, i.e., the target components, the background components, and the noise components [51]. We have

$$f_D(x, y) = f_T(x, y) + f_B(x, y) + f_N(x, y), \quad (1)$$

where f_D, f_T, f_B , and f_N are the original infrared image, the target image, the background image, and the noise image, respectively. (x, y) denotes the pixel coordinates. Gao et al. [19] decomposed the input image through image patch level, defined as

$$D = T + B + N, \quad (2)$$

where D, T, B , and N are the original patch image, target patch image, background patch image, and random noise patch, respectively. These patch images can be obtained by sliding windows from the top left to the bottom right on the image. Because the background is considered to be slowly transitioning, the adjacent background patches are approximately linearly related, whereas small targets generally contain few pixels (1×1 to 9×9). Therefore, the background image can be modeled as a low-rank component, and the target image is usually considered to be a sparse component. Thus, the small-target detection problem can be formulated as a robust principal component analysis model (RPCA), such as in [26,52,53]. However, the disadvantage of the IPI model is that it destroys the local features between pixels. To solve this problem, the researchers propose the IPT model, as

$$\mathcal{D} = \mathcal{T} + \mathcal{B} + \mathcal{N}, \quad (3)$$

where $\mathcal{D}, \mathcal{T}, \mathcal{B}$ and \mathcal{N} are the original patch tensor, target patch tensor, background patch tensor, and noise patch tensor, respectively. Within an image, the patch image can be obtained by expanding the patch tensor through mode-3 [43]. Thus, the IPT model can utilize more spatial information than the IPI model. Generally, the sparse target tensor satisfies $\|\mathcal{T}\|_0 < \tau$, where τ is a constant determined by target size. Thus, without considering noise, we can formulate a tensor PCA model that separates the sparse target from the background, as

$$\min_{\mathcal{B}, \mathcal{T}} \text{rank}(\mathcal{B}) + \lambda \|\mathcal{T}\|_0 \quad \text{s.t. } \mathcal{D} = \mathcal{B} + \mathcal{T}, \quad (4)$$

where $\|\cdot\|_0$ is the l_0 norm, and λ is the compromise parameter.

3.2. Laplace-Based Rank Approximation

As in previous analyses, how to approximate the background rank is a key issue for constructing the IPT model. Previously, researchers used TNN [45] as a convex relaxation of multirank tensor's l_1 norm. However, TNN accredits the equivalent weight to

each singular value, so it may falsely separate some background clutter as the target, which will result in false alarms. Therefore, using TNN directly in the IPT model will reduce the detection performance. Recently, the Laplacian function [54] was introduced to TNN to generate multirank nonconvex approximations of tensors for solving low-rank tensor completion problems, and this approach is defined as follows,

$$\begin{aligned} \|\mathcal{X}\|_{LAP-TNN} &= \sum_{k=1}^{n_3} \sum_{i=1}^n \phi\left(\sigma_i\left(\overline{\mathcal{X}}^{(k)}\right)\right) \\ &= \sum_{k=1}^{n_3} \sum_{i=1}^n \left(1 - e^{-\sigma_i\left(\overline{\mathcal{X}}^{(k)}\right)/\epsilon}\right), \end{aligned} \tag{5}$$

where $\phi(x) = \left(1 - e^{-x/\epsilon}\right)$, its bandwidth is determined by the positive parameter ϵ , and $n = \min(n_1, n_2)$. $\sigma_i(O)$ is the i th singular value of matrix O . As shown in Figure 1, the Laplace function can better approximate the L_0 norm than the L_1 norm. Therefore, we introduce the LAP-TNN to constrain the background tensor, and our model is formulated as

$$\min_{\mathcal{B}, \mathcal{T}} \|\mathcal{B}\|_{LAP-TNN} + \lambda \|\mathcal{T}\|_0 \quad s.t. \mathcal{D} = \mathcal{B} + \mathcal{T}. \tag{6}$$

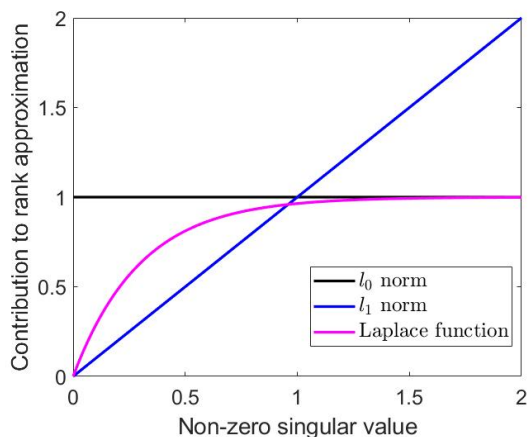


Figure 1. Comparison of l_0 norm, l_1 norm, and approximate rank of Laplace function.

3.3. Construction of the Local Prior

RPCA cannot preserve edge components well in the background image. Researchers often formulate some prior information to preserve edge components and to better estimate targets. The RIPT model attempts to retain the small targets while suppressing strong edges by using the local structural information [27]. However, this weight cannot perceive the corner structure in the background, resulting in false alarms and target oversuppression [55]. The PSTNN model further analyses the structural tensor map and proposes an improved local structural weight map. However, the performance of this local structural weight map in suppressing edges is also unsatisfactory. In this paper, we design a novel target prior with the following feature information.

A. Gradient feature information

In general, small targets in local areas usually cause gray change, and the gradient vectors in the surrounding areas basically have a tendency to target center. Therefore, the local gradient information can be beneficial to detect small targets. LIG [14] introduces gradient information that can further improve detection performance. LIG divides an image patch into four parts ($\Phi_i, i = 1, \dots, 4$) and calculates the gradient mean square of each small patch separately. We have

$$G_i = \frac{1}{N_i} \sum_{k=1}^{N_i} \|g_{\Phi_i}(k)\|^2, (i = 1, \dots, 4), \tag{7}$$

where g_{Φ_i} is the gradient magnitude of which element gradient pointed to the center in Φ_i and N_i is the element number that satisfies the direction constraint in the part Φ_i . The maximum and minimum values of G_i can be obtained by

$$G_{max} = \max\{G_i\}, \quad G_{min} = \min\{G_i\}, \quad (i = 1, \dots, 4). \tag{8}$$

Finally, the gradient diagram can be calculated by

$$G = \begin{cases} \sum_{i=1}^4 G_i, & \text{if } \frac{G_{min}}{G_{max}} > k_1 \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

in which G is the gradient intensity of the image patch, and k_1 is a constant to reduce some clutter and noise.

In this paper, we propose multidirectional LIG to better measure the gradient strength of the elements in the patch. We divide an image patch into four quadrants and calculate the gradient in eight directions, as shown in Figure 2.

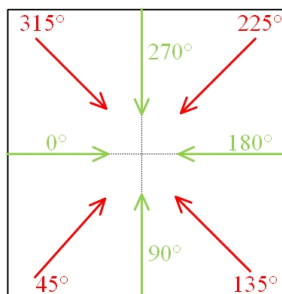


Figure 2. Gradient direction map.

We set the 45°, 135°, 225°, and 315° as the primary directions and 0°, 90°, 180°, and 270° as the secondary directions. The sum of the gradient mean square values of which elements gradient along the primary and secondary directions is calculated for each quadrant as $Q_i, (i = 1, 2, 3, 4)$, and different weights are assigned to the primary and secondary directions.

$$G_{i,D} = \frac{1}{N} \sum_{j=1}^N \|g_{\Phi_{i,D}}(j)\|^2 \quad (D = 0^\circ, 45^\circ, \dots, 315^\circ) \tag{10}$$

$$\begin{aligned} Q_1 &= 0.7 * G_{1,225^\circ} + 0.15 * G_{1,270^\circ} + 0.15 * G_{1,180^\circ} \\ Q_2 &= 0.7 * G_{2,315^\circ} + 0.15 * G_{2,270^\circ} + 0.15 * G_{2,0^\circ} \\ Q_3 &= 0.7 * G_{3,45^\circ} + 0.15 * G_{3,0^\circ} + 0.15 * G_{3,90^\circ} \\ Q_4 &= 0.7 * G_{4,135^\circ} + 0.15 * G_{4,90^\circ} + 0.15 * G_{4,180^\circ} \end{aligned} \tag{11}$$

where $g_{\Phi_{i,D}}$ is the gradient magnitude of which element gradient along the direction D in part Φ_i . N is the element number that satisfy the direction constraint in the part Φ_i . The maximum and minimum values of Q can be calculated by

$$Q_{max} = \max\{Q_i\}, \quad Q_{min} = \min\{Q_i\} \quad (i = 1, \dots, 4). \tag{12}$$

Finally, the multidirectional gradient diagram can be calculated by

$$Q_g = \begin{cases} \sum_{i=1}^4 Q_i, & \text{if } \frac{Q_{min}}{Q_{max}} > k_2 \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where Q_g is the comprehensive gradient of image patch, and k_2 is a parameter to reduce some clutter and noise. In this paper, k_2 is set to 0.5.

B. Edge and highlighted area indicator

The target is enhanced by the gradient information in (13), and the highlighted areas and edges are also enhanced. To eliminate strong edge clutter, we introduce highlighted areas and edge retain weight. Li et al. [56] proposed an edge weight formula in the following form,

$$E_G(q) = \frac{1}{N} \sum_{p=1}^N \frac{\theta_{G_1}^2(q) + \delta}{\theta_{G_1}^2(p) + \delta'} \quad (14)$$

where G is the guide image, $\theta_{G_1}^2(q)$ is the variance of G for a 3×3 window centred on q , δ is a small constant that has a value of $(0.001 \times R)^2$, R is the dynamic range of the input image, and N is the total pixel number in an image. This edge weight can perceive sharp edges, but we expect to measure the highlighted areas too, so we made the following improvements. We modify the 3×3 window variance with the 5×5 window mean to obtain an edge and highlighted areas indicator with the following equation,

$$EH_G(q) = \frac{1}{N} \sum_{p=1}^N \frac{M_{G_2^2}(q) + \delta}{M_{G_2^2}(p) + \delta'} \quad (15)$$

where G^2 is the square of the input image and $M_{G_2^2}(p)$ is the mean of a 5×5 window centred on pixel p in G^2 . Edge and highlighted area-retained weighting maps can be obtained by filtering the original map by using (15):

$$I_{EH} = EH_G(f_D(x, y)). \quad (16)$$

C. Local prior calculation

The gradient feature enhances the target, but the highlighted areas and edges are also enhanced. The edge and highlighted area indicator measures the highlighted areas and edge areas of the original image. Therefore, we combine the gradient feature information with the edge and highlighted area indicator map to calculate the local target likelihood,

$$W_t = \max(Q_g - \beta \cdot I_{EH}, 0), \quad (17)$$

where β is a regularization factor. To normalize the range of W_t , we linearly stretch it as follows,

$$W_t = \frac{W_t - W_{min}}{W_{max} - W_{min}}, \quad (18)$$

where W_{max} and W_{min} represent the maximum and minimum values in W_t , respectively. One example of our local target prior weight is shown in Figure 3.

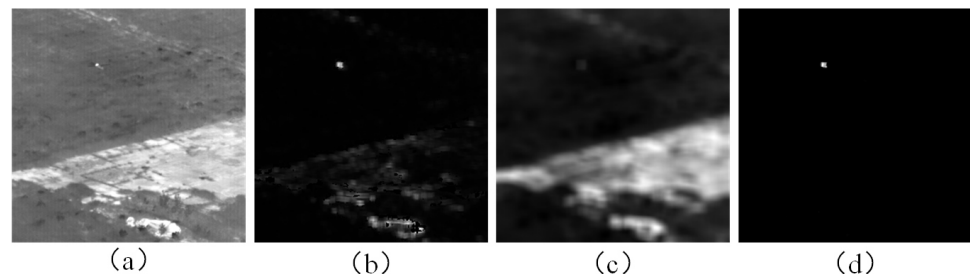


Figure 3. Small-target local priori weight. (a) Input image. (b) Gradient features. (c) Edge and highlighted area indicator. (d) Target local prior.

3.4. The Proposed Enhanced IPT Model

Introducing the local prior to IPT model, it can be updated as

$$\min_{\mathcal{B}, \mathcal{T}} \|\mathcal{B}\|_{LAP-TNN} + \lambda \|\mathcal{W}_{re} \odot \mathcal{T}\|_1 \quad s.t. \mathcal{D} = \mathcal{B} + \mathcal{T}, \quad (19)$$

where \odot is the Hadamard product. \mathcal{W}_{re} is the reciprocal of \mathcal{W}_t . To prevent \mathcal{W}_{re} from going to infinity, we add a small positive number τ to \mathcal{W}_t .

$$\mathcal{W}_{re} = \frac{1}{\tau + \mathcal{W}_t} \quad (20)$$

Many researchers [57–59] have employed reweighting methods to speed up the convergence and reduce the iteration calculation. We adopted the following reweighting method to speed up the proposed algorithm iteration process,

$$\mathcal{W}_{se}^{k+1} = \frac{c}{|\mathcal{T}|^k + \eta}, \quad (21)$$

where c is a positive constant, $k + 1$ denotes $(k + 1)$ iterations, and η is a small positive number that avoids dividing by 0. In general, c is set to 1 [29,60]. Then, we combine the two indicators to obtain the final target likelihood, as follows:

$$\mathcal{W} = \mathcal{W}_{se} \odot \mathcal{W}_{re}. \quad (22)$$

Finally, the proposed IPT model in (19) is updated as

$$\min_{\mathcal{T}, \mathcal{B}} \|\mathcal{B}\|_{LAP-TNN} + \lambda \|\mathcal{W} \odot \mathcal{T}\|_1 \quad s.t. \mathcal{D} = \mathcal{T} + \mathcal{B}. \quad (23)$$

3.5. Model Solution

The ADMM [58] is a common technology for solving the optimization problem with constraints, which is characterized by fast convergence and high accuracy. In this section, we employ ADMM to solve the proposed model. The augmented generalized Lagrangian function for Model (23) is formulated as

$$L(\mathcal{B}, \mathcal{T}, \mathcal{W}, \mathcal{Z}) = \|\mathcal{B}\|_{LAP-TNN} + \lambda \|\mathcal{W} \odot \mathcal{T}\|_1 + \langle \mathcal{Z}, \mathcal{T} + \mathcal{B} - \mathcal{D} \rangle + \frac{\mu}{2} \|\mathcal{T} + \mathcal{B} - \mathcal{D}\|_F^2, \quad (24)$$

where \mathcal{Z} is the Lagrange multiplier, $\langle \cdot \rangle$ is the inner product of two tensors, and μ is a penalty factor greater than 0. The unknown variables in (24) are separated, and we can solve them in an alternate manner. Problem (24) can be decomposed into two subproblems:

$$\mathcal{T}^{k+1} = \arg \min_{\mathcal{T}} \lambda \|\mathcal{W}^k \odot \mathcal{T}\|_1 + \frac{\mu^k}{2} \|\mathcal{T} + \mathcal{B}^k - \mathcal{D} + \frac{\mathcal{Z}^k}{\mu^k}\|_F^2 \quad (25)$$

$$\mathcal{B}^{k+1} = \arg \min_{\mathcal{B}} \|\mathcal{B}\|_{LAP-TNN} + \frac{\mu^k}{2} \|\mathcal{T}^{k+1} + \mathcal{B} - \mathcal{D} + \frac{\mathcal{Z}^k}{\mu^k}\|_F^2. \quad (26)$$

(1) We use the soft thresholding operator [61] to solve subproblem (25), and its solution can be given by

$$\mathcal{T}^{k+1} = \text{Softshrink}(\mathcal{D} - \mathcal{B}^k - \frac{\mathcal{Z}^k}{\mu^k}, \frac{\lambda \mathcal{W}^k}{\mu^k}). \quad (27)$$

(2) To solve the low-rank tensor \mathcal{B} , subproblem (26) can be expressed in the form

$$\arg \min_{\mathcal{B}} \|\mathcal{B}\|_{LAP-TNN} + \frac{\mu}{2} \|\mathcal{B} - \mathcal{H}\|_F^2, \quad (28)$$

where $\mathcal{B}, \mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, $\mathcal{H} = \mathcal{D} - \mathcal{T}^{k+1} - \frac{\mathcal{Z}^k}{\mu^k}$. To simplify the expression, we omit the iteration k in (28). From Equation (5), we can observe that the nonconvex tensor rank surrogate is a linear combination of all frontal slice Laplace functions in the Fourier domain along the tube dimension [29,62]. Therefore, the optimization problem in (28) can be transformed into n_3 optimization problems in the Fourier domain [54], as follows,

$$\arg \min_{\overline{\mathcal{B}}^{(s)}} \sum_{i=1}^n \phi(\sigma_i(\overline{\mathcal{B}}^{(s)})) + \frac{\mu}{2} \|\overline{\mathcal{B}}^{(s)} - \overline{\mathcal{H}}^{(s)}\|_F^2, \quad (29)$$

where $s = 1, \dots, n_3$ and $\overline{H}^{(k)}$ and $\overline{B}^{(s)} \in \mathbb{C}^{n_1 \times n_2}$. The generalized weighted singular value thresholding operator [63] can solve model (29). We have

$$\overline{B}^{(s)} = \overline{U}^{(s)} \overline{S}^{(s)}_{\frac{\nabla\phi}{\mu}} \overline{V}^{(s)} \tag{30}$$

in which $\overline{H}^{(s)} = \overline{U}^{(s)} \overline{S}^{(s)} \overline{V}^{(s)H}$, and

$$\overline{S}^{(s)}_{\frac{\nabla\phi}{\mu}}(i, i) = \max \left\{ \left(\overline{S}^{(s)}(i, i) - \frac{\nabla\phi(\sigma_i^s)}{\mu} \right), 0 \right\}, \tag{31}$$

where $\nabla\phi(\sigma_i^s) = 1/\epsilon \times \exp(-\sigma_i^s/\epsilon)$ is the gradient of σ at σ_i^s and σ_i^s is the i -th singular value of $\overline{B}^{(s)}$. \mathcal{B} can be obtained by inverse FFT. Algorithm 1 presents the solution procedure for model (28).

Algorithm 1 Solution process for model (28)

Input: $\overline{B}^k, \mathcal{H}, \mu^k, \epsilon$

Output: $\overline{B}^{k+1}, \mathcal{B}^{k+1}$

Step 1. Compute $\overline{H} = \text{fft}(\mathcal{H}, [], 3)$

Step 2. Compute each forward slice of \overline{B}^{k+1} as:

for $s = 1, \dots, \lceil (n_3 + 1)/2 \rceil$ **do**

(1) $[\overline{U}^{(s)}, \overline{S}^{(s)}, \overline{V}^{(s)}] = \text{SVD}(\overline{H}^{(s)})$.

(2) Calculate $\overline{S}^{(s)}_{\frac{\nabla\phi}{\mu^k}}$ by Equation (30)

(3) $(\overline{B}^{k+1})^{(s)} = \overline{U}^{(s)} * \overline{S}^{(s)}_{\frac{\nabla\phi}{\mu^k}} * \overline{V}^{(s)H}$

end for

for $s = \lceil (n_3 + 1)/2 \rceil + 1, \dots, 1$ **do**

$(\overline{B}^{k+1})^{(s)} = \text{conj}((\overline{B}^{k+1})^{(n_3-s+2)})$

end for

Step 3. Compute $\mathcal{B}^{k+1} = \text{ifft}(\overline{B}^{k+1}, [], 3)$

\mathcal{Z} and μ are updated as follows,

$$\mathcal{Z}^{k+1} = \mathcal{Z}^k + \mu^k (\mathcal{D} - \mathcal{T}^{k+1} - \mathcal{B}^{k+1}) \tag{32}$$

$$\mu^{k+1} = \rho \mu^k, \tag{33}$$

where ρ is a positive constant. The proposed model's solution process is shown in Algorithm 2.

Algorithm 2 The proposed model solved by ADMM**Input:** $\mathcal{D}, \mathcal{W}, \lambda, \mu_0, \epsilon$ **Output:** $\mathcal{B}^k, \mathcal{T}^k$ Initialization: $\mathcal{B}^0 = \mathcal{T}^0 = \mathcal{Z}^0 = 0, \mathcal{W}_{sw} = 1, \mathcal{W}^0 = \mathcal{W}_{sw} \odot \mathcal{W}_{re}, \mu^0 = 6 \times 10^{-3}, \rho = 1.05, k = 0, tol = 1 \times 10^{-3}$ **while** $\frac{\|\mathcal{T}^k + \mathcal{B}^k - \mathcal{D}\|_F}{\|\mathcal{D}\|_F} > tol$ and $\|\mathcal{T}^{k+1}\|_0 \neq \|\mathcal{T}^k\|_0$ **do** update \mathcal{T}^{k+1} by (27); update \mathcal{B}^{k+1} by Algorithm 1 update \mathcal{W}^{k+1} by (22); update \mathcal{Z}^{k+1} by (32); update μ^{k+1} by (33); update $k = k + 1$.**end while**

3.6. Whole Process of the Proposed Method

Figure 4 illustrates the whole process of the proposed method, as follows.

- (1) Local target prior calculation. For an obtained infrared image, we calculate its gradient feature by (13) and edge and highlighted area indicator by (16), and combine them into the local target prior weight \mathcal{W}_t by (18).
- (2) Patch tensor formulation. Sliding a window of size $s \times s$ over original image from top left to bottom right, we stack the paths into the original image 3D tensor $\mathcal{D} \in \mathbb{R}^{s \times s \times z}$. Similarly, the prior weight tensor $\mathcal{W}_t \in \mathbb{R}^{s \times s \times z}$ can be constructed.
- (3) The input tensor is decomposed into a background tensor \mathcal{B} and a target tensor \mathcal{T} by using ADMM in Algorithm 2.
- (4) The 2D background image and target image can be calculated from the background tensor \mathcal{B} and target tensor \mathcal{T} by employing 1D median filter on the overlapping positions.

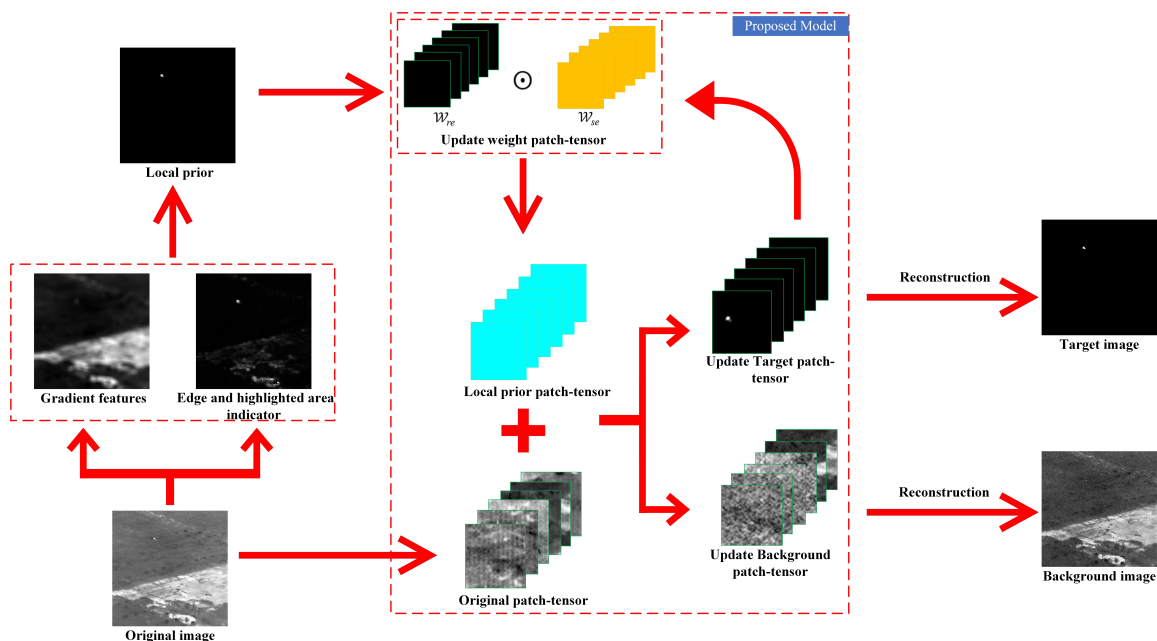


Figure 4. Procedure of the proposed algorithm.

4. Experiments and Results

In this section, we validate the proposed model's performance in terms of background suppression, target enhancement, and target detection abilities. Seven state-of-the-art methods are included for comparison, including WLDM [64], Top-hat [7], IPI [19], SMSL [25], RIPT [27], PSTNN [28], and LogTFNN [43]. WLDM and Top-hat are local contrast-based

methods, IPI and SMSL are IPI model-based methods, and RIPT, PSTNN, and LogTFNN are similar to our model and are IPT model-based methods.

4.1. Evaluation Metrics

We use a variety of metrics to measure the performance of the comparison detection methods. The signal-to-clutter ratio gain (SCRG) is a common index that measures the saliency of the target, and a higher SCRG value indicates that the detection model could enhance the target more distinctly. It is related to the SCR, which is defined as

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b}, \quad (34)$$

where μ_t and μ_b denote the average value of the target pixels and neighboring background regions, respectively. σ_b denotes the variance in the pixel values in the background neighbourhood. SCRG is then defined as

$$SCRG = \frac{SCR_{out}}{SCR_{in}}, \quad (35)$$

where SCR_{in} represents the SCR of the original image and SCR_{out} represents the SCR of the target image.

The background suppression factor (BSF) evaluates the background clutter suppressing ability of one detection method; the higher the BSF value is, the more effectively the background clutter is suppressed. We have

$$BSF = \frac{\sigma_{in}}{\sigma_{out}}. \quad (36)$$

In addition, we employ ROC curves to comprehensively evaluate the detection performance. In a ROC curve, the vertical axis represents the model's true positive rate (TPR), and the horizontal axis represents model's false-positive rate (FPR):

$$TPR = \frac{\text{number of detected true targets}}{\text{number of actual targets}} \quad (37)$$

$$FPR = \frac{\text{number of false pixels detected}}{\text{number of pixels in whole image}}. \quad (38)$$

4.2. Parameters Analysis

In our model, there are some prominent parameters that influence the detection performance of the proposed model, such as the patch size, sliding step size, tradeoff parameter λ , penalty coefficient μ , and regularization factor β . To better analyse the role of each parameter, we conduct parametric experiments by varying one variable while fixing the others. Figure 5 depicts the ROC curves for four real captured IR sequences with various parameter settings. It should be noted that the parameter settings presented in this section are not optimal.

The patch size plays a crucial role in our model. A large patch size usually leads to poor detection performance because the model may identify some sparse noise as the target. Then, a small patch size would weaken the sparsity of the target. For the four test sequences, the ROC curves with regard to the patch size that range from 20 to 60 with an interval of 10 are displayed in Figure 5a. It seems that the best detection performance can be obtained when the patch size is set to 50.

Another key parameter is the sliding step. A larger sliding step could decrease the slicing number of the IPT, which can destroy the nonlocal correlation of the background. In contrast, a smaller sliding step could slow down the construction of the tensor, and target sparsity would decrease as the number of target slices increases. We change the

step size from 20 to 60 at intervals of 10, and the best detection performance can be reached when the sliding step is set to 40, as shown in Figure 5b.

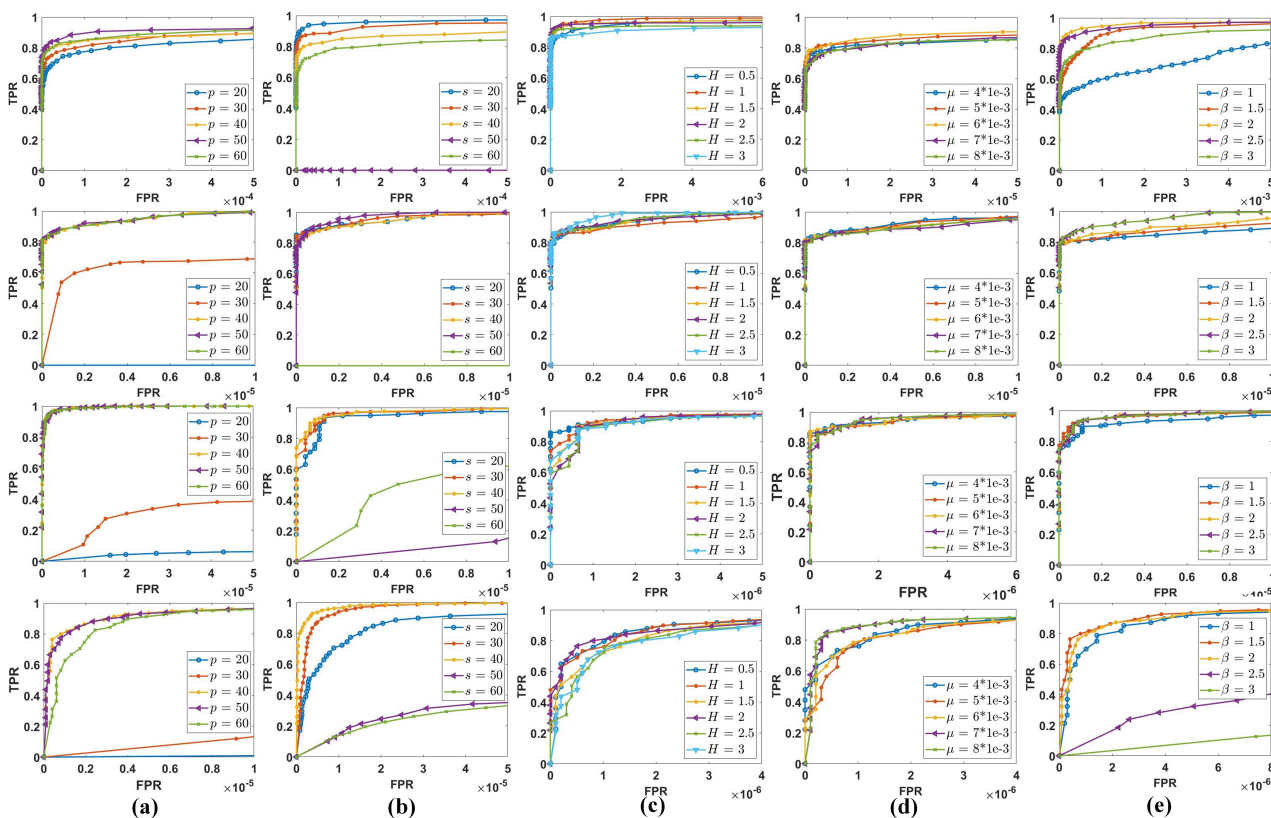


Figure 5. ROC curves are formed with different parameters. (a) The first column depicts the curves of different patch sizes p . (b) The second column depicts the output of different sliding steps s . (c) The third column shows the influence of the tradeoff factor λ . (d) The fourth column shows the results of different penalty factors μ . (e) The fifth column shows the result for different β values. Each row represents a different parameter in the same sequence.

The weight parameter λ is used to balance the target’s sparsity and the background’s low-rank. When λ is large, a heavy penalty acts on the sparse components, and may generate an incomplete target; when λ is small, some false alarm might emerge with the true target. Referring to [28], we set λ to $\frac{H}{\sqrt{\max(n_1, n_2) \times n_3}}$ and then adjust H in the range [0.5, 3] instead of λ . Figure 5c shows the effect of H on the four sequences, and we can easily observe that the performance is relatively good when $H = 2$. Thus, in the experiment, we set λ as $\lambda = \frac{2}{\sqrt{\max(n_1, n_2) \times n_3}}$.

The penalty factor μ also plays a significant role in separating the small-target component and the low-rank background. It also influences the optimization process and the convergence speed. A smaller μ will cause certain target components to be preserved in the background components, which may shrink the target component. If μ is large, although the target component can be perfectly retained, some background component may appear in the target patch tensor, too. We varied μ in the four sequences from 0.004 to 0.008 at 0.001 intervals, as shown in Figure 5d. The best detection performance is reached when $\mu = 0.006$.

The regularisation factor β is employed to eliminate the clutter that is generated by the local gradient calculation. A large β can remove background clutter, but the target may be shrunk; a small β could retain a full target, but some background clutter may be also retained. We varied β from 1 to 3 at intervals of 0.5, and the best noise rejection performance is obtained when $\beta = 2$, as shown in Figure 5e.

4.3. Comparative Evaluation

We tested the robustness of the proposed model by conducting comparison experiments on a sequence of six different scenarios. Table 2 describes the details of the six real sequences. These real infrared sequences are selected from the dataset provided by the authors in reference [65]. We choose seven advanced methods to compare with the proposed method, and Table 3 gives the details of the parameters for the comparison methods. The visual output on sequences 1 to 6 of different methods are shown in Figure 6. As shown in Figure 6, the Top-hat and WLCM methods enhance the target to some extent, but the strong interference part cannot be suppressed, and the detection results are poor. This is because they use a fixed filter structure to process the image, which makes them more sensitive to the background and leads to a less robust final result. Compared with the first two algorithms, IPI improved detection performance, but the detection results still retain some clutter due to the inability to accurately approximate the background rank. RIPT, PSTNN, and LogTFNN utilize image information in tensor form to separate targets more effectively than IPI, but these algorithms are unable to accurately calculate the background rank, resulting in lower robustness to different backgrounds. Compared to these methods, the proposed method can more accurately approximate the background rank by using the Laplace function. The constructed target prior weights can suppress strong edges while enhancing the target. As a result, the proposed method can better separate the background and target components than other methods.

Table 2. Sequence information for six different scenes.

	Frames	Length	Target and Background Description
Seq.1	100	256 × 256	Single target, ground background
Seq.2	100	256 × 256	Single target, open space background
Seq.3	150	256 × 256	Single target, open space background
Seq.4	150	256 × 256	Single target, open space background
Seq.5	70	256 × 256	Single target, ground background
Seq.6	150	256 × 256	Single target, ground background

Table 3. Parameters of the eight comparison methods.

Methods	Parameter Settings
Top-hat [7]	Structure pattern: square, size 3 × 3
WLCM [64]	Neighbourhood structure: (3 ~ 5) × (3 ~ 5)
IPI [19]	Patch size: 50 × 50, sliding step: 10, $\lambda = \frac{1}{\sqrt{(\min(m,n))}}$, $\epsilon = 10^{-7}$
SMSL [25]	Patch size: 30 × 30, sliding step: 30, $\lambda = \frac{1}{\sqrt{(\min(m,n))}}$, $\epsilon = 10^{-7}$
RIPT [27]	Patch size: 30 × 30, sliding step: 10, $\lambda = \frac{1}{\sqrt{(\min(m,n))}}$, $\epsilon = 10^{-7}$, $h = 1$
PSTNN [28]	Patch size: 40 × 40, sliding step: 40, $\lambda = \frac{2}{\sqrt{(\min(m,n))}}$, $\epsilon = 10^{-7}$
LogTFNN [43]	Patch size: 40 × 40, sliding step: 40, $\lambda = \frac{0.4}{\sqrt{(\max(n_1, n_2) \times n_3)}}$, $\beta = 0.05, \mu = 200$
Proposed	Patch size: 50 × 50, sliding step: 40, $\lambda = \frac{2}{\sqrt{(\max(n_1, n_2) \times n_3)}}$, $\beta = 2, \mu = 0.006, tol = 10^{-3}$

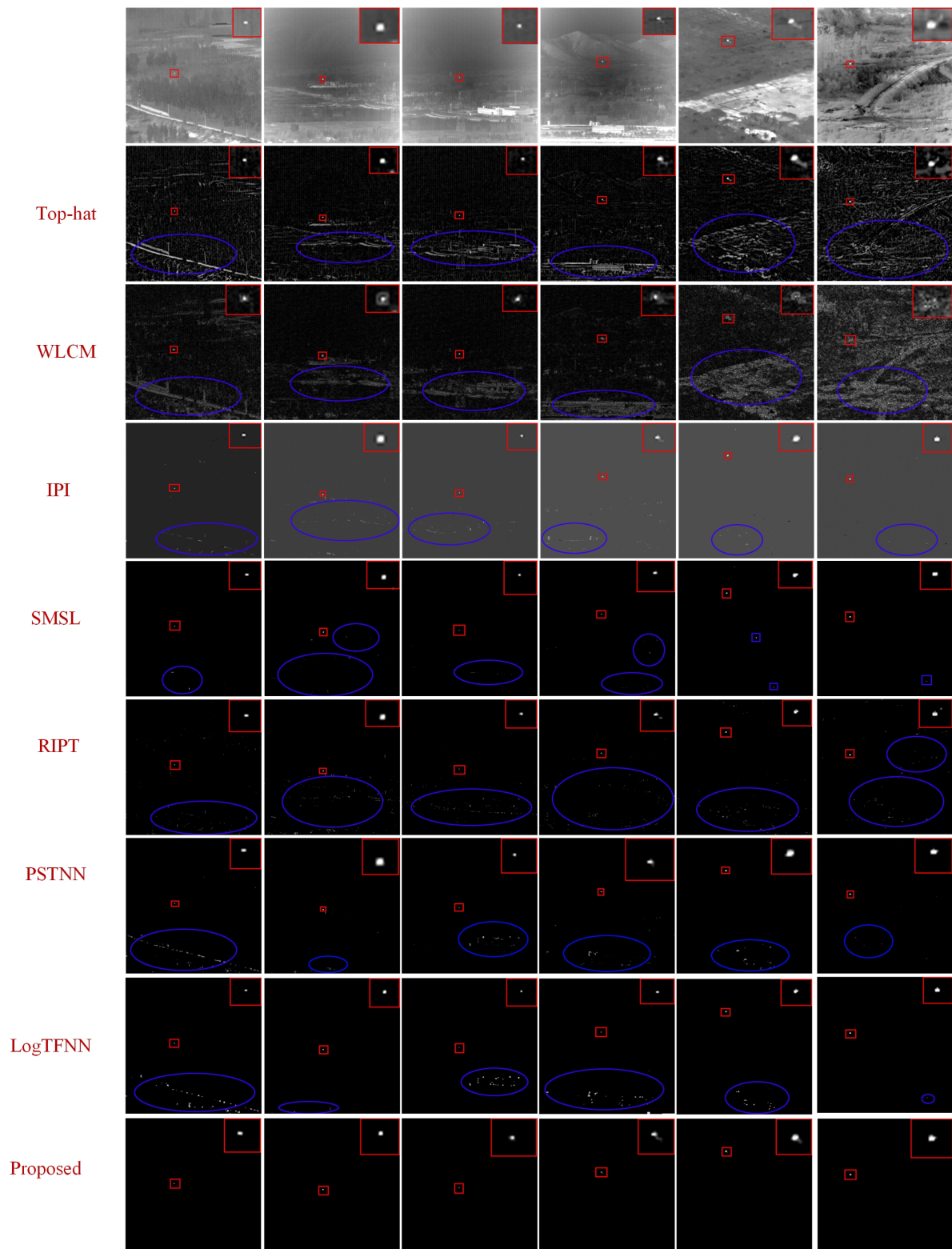


Figure 6. Results of the 8 methods on the 6 sequences.

Table 4 displays the BSF and SCRG metrics for the six sequences by the eight methods, wherein the best result for each metric is highlighted in red and the next best result is marked in green. Larger BSF and SCRG values indicate better detection performance.

As displayed in Table 4, the proposed method achieves more significant advantages in SCRG and BSF. PSTNN achieves the next best results on Sequence 2 and Sequence 4, which indicates that the tensor-based model can indeed improve the detection performance

by obtaining more spatial information. In addition, nonlocal prior methods can achieve higher values than local prior methods (Top-hat, WLCM). This is due to the fact that nonlocal prior-based methods can improve detection performance by utilizing additional spatial information.

Table 4. The BSF and SCRG results for the six sequences by eight methods.

Method	Seq.1		Seq.2		Seq.3		Seq.4		Seq.5		Seq.6	
	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG
Top-hat	2.62	2.28	4.67	5.16	4.82	3.36	4.84	7.46	5.08	4.53	4.17	3.82
WLCM	5.93	8.76	8.16	18.22	8.20	9.15	9.20	14.20	10.12	3.80	8.15	2.85
IPI	11.10	4.59	12.02	27.00	15.65	13.43	13.05	17.97	17.66	14.28	17.12	17.85
SMSL	31.64	18.50	8.50	30.17	55.43	18.23	37.68	27.69	82.02	14.84	115.86	18.81
RIPT	10.21	13.26	8.71	19.91	14.29	10.61	13.74	17.49	14.82	10.08	11.90	10.31
PSTNN	6.37	9.25	13.20	30.19	64.66	17.38	47.97	28.16	32.87	13.52	34.67	18.16
LogTFNN	3.35	3.65	9.54	20.59	6.17	3.89	9.33	8.97	106.78	12.60	97.86	16.06
Proposed	75.63	29.89	59.57	53.18	269.76	25.14	371.86	46.19	327.44	20.11	107.98	22.75

* For each index, the best value is colored with red, and the second best value is colored with green.

To further demonstrate the comprehensive performance of the compared methods, we show the ROC curves of the different methods for six sequences (Figure 7). The ROC curves show that the proposed method works best, the PSTNN, RIPT, and SMSL methods have sufficient ROC performance, whereas the other methods have typical performance. Overall, the proposed method has the highest probability of detection for the same false alarm rate and the lowest false alarm rate for the same accuracy on the different scenes.

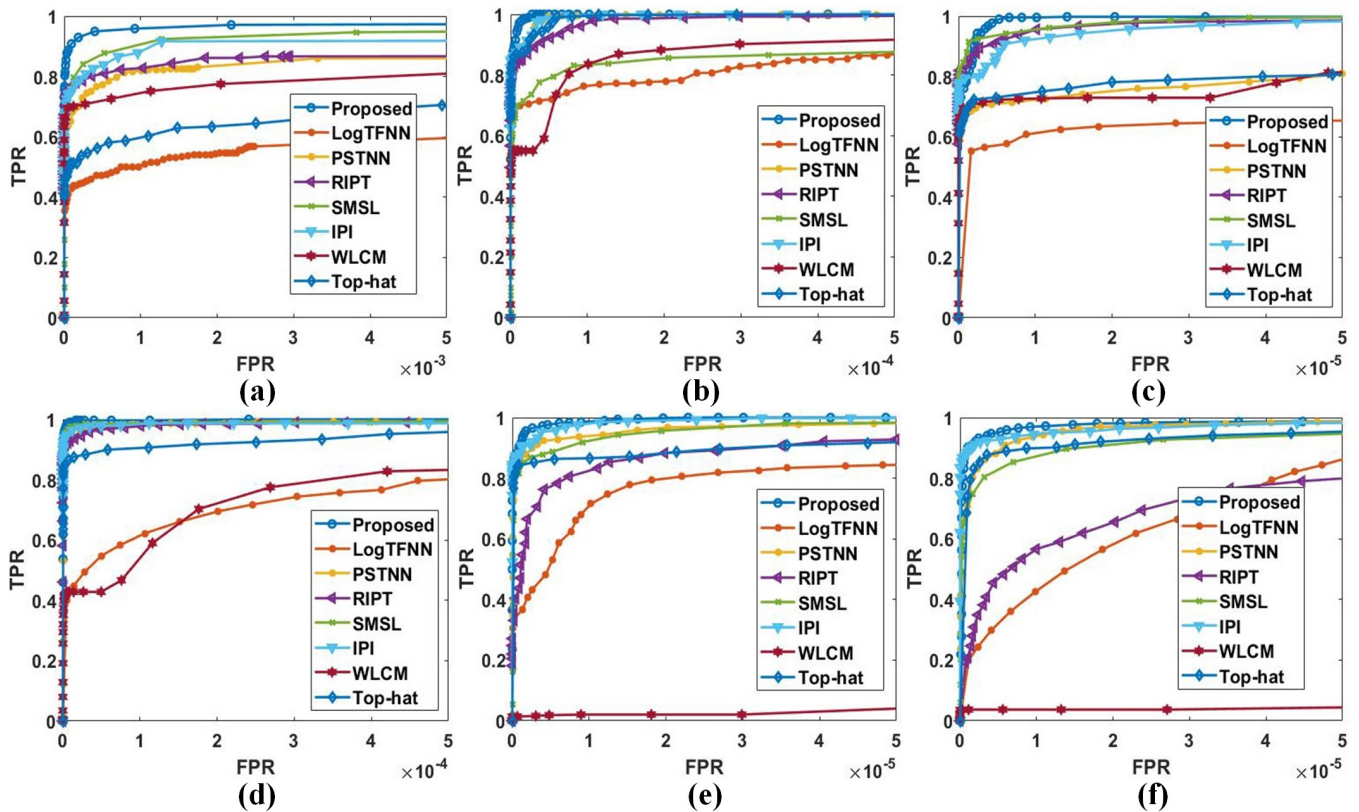


Figure 7. ROC curves for six sequences by comparison methods. (a) Sequence 1. (b) Sequence 2. (c) Sequence 3. (d) Sequence 4. (e) Sequence 5. (f) Sequence 6.

4.4. Robustness in Different Scenes

The detection performance of a method is evaluated based on its accuracy and robustness in different scenarios, namely, scenes with diverse backgrounds, target variability, and uncertain target locations. As the distance between a target and the infrared sensor is variable, the target pixel size in an image and the target brightness are also variable. To further verify the robustness of each algorithm, we selected 20 small target images with different backgrounds and different intensities from the public dataset in [66,67] to test their performance, as shown in Figure 8.

The target image output of the proposed method on 20 single-frame images are displayed in Figure 9, and their corresponding 3D maps of the target images are displayed in Figure 10. We can see that various types of background clutter are completely removed in Figure 9. In Figure 10, the background components are suppressed to a large extent, and the target becomes the most prominent elements in the 3D map. As shown in Figures 9 and 10, all small targets are significantly enhanced by the proposed method, and various types of background clutter are effectively suppressed. From Table 4, we can observe that PSTNN, which is also based on T-SVD, has a better BSF and SCRG metric than RIPT and LogTFNN on most sequences. Therefore, we choose PSTNN as a visual comparison. Figure 11 shows the PSTNN results of 20 single-frame images, and the corresponding 3D maps are shown in Figure 12. As we can see, although PSTNN could detect the target out for the 20 images, there is still some background clutter (marked by blue boxes) in the output image. Therefore, by comparison, our method could detect more real target components in different scenes than the baseline methods.

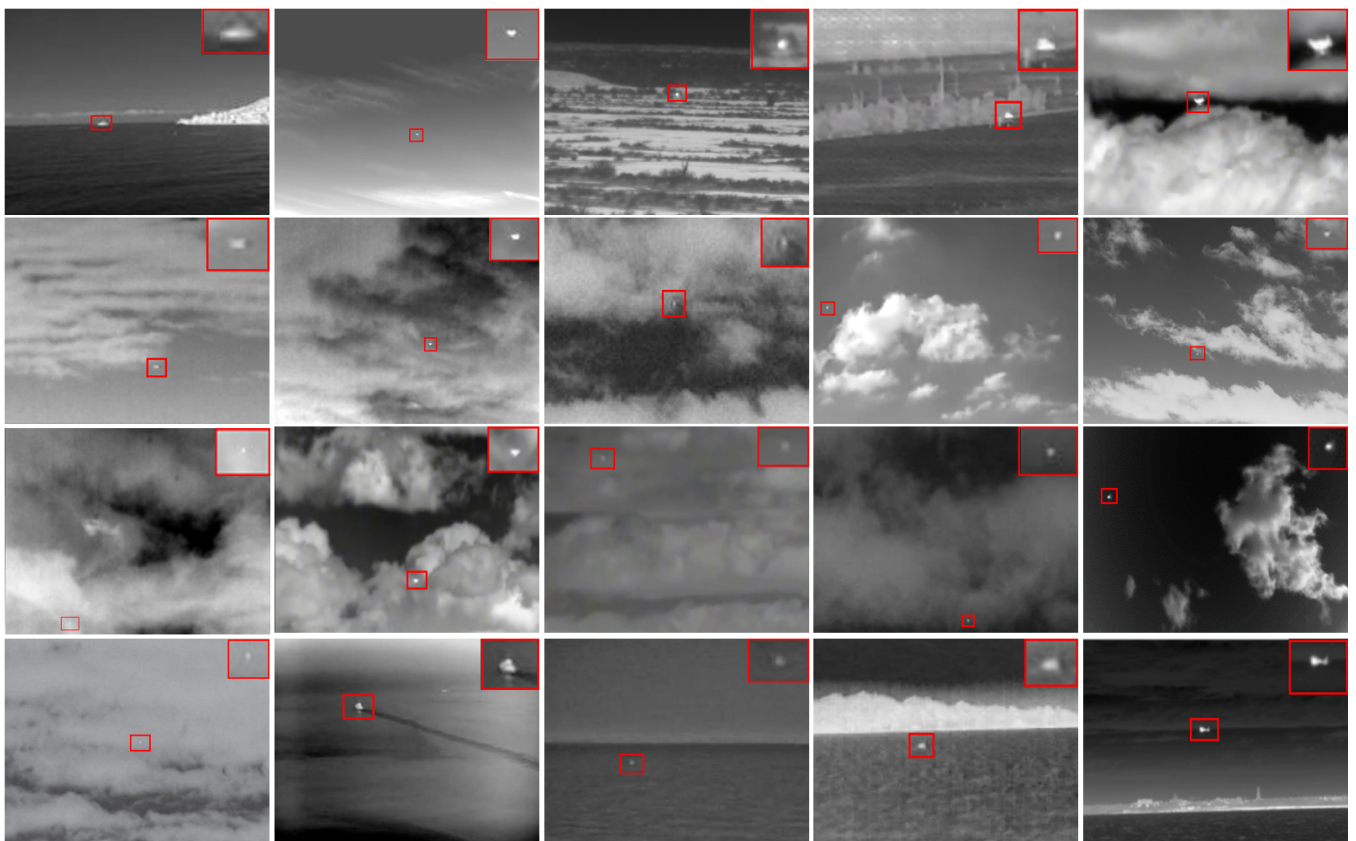


Figure 8. Twenty single-frame images in different scenes.

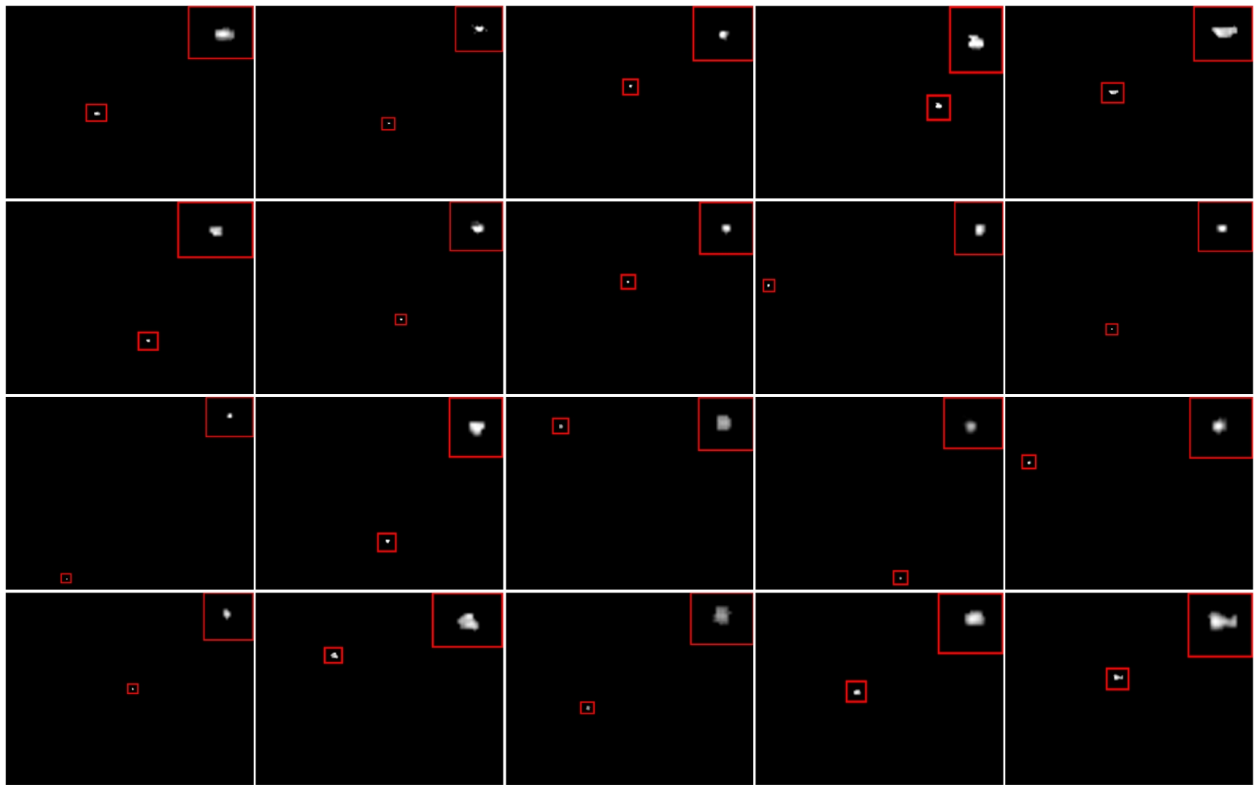


Figure 9. Process output of proposed method for 20 single-frame images.

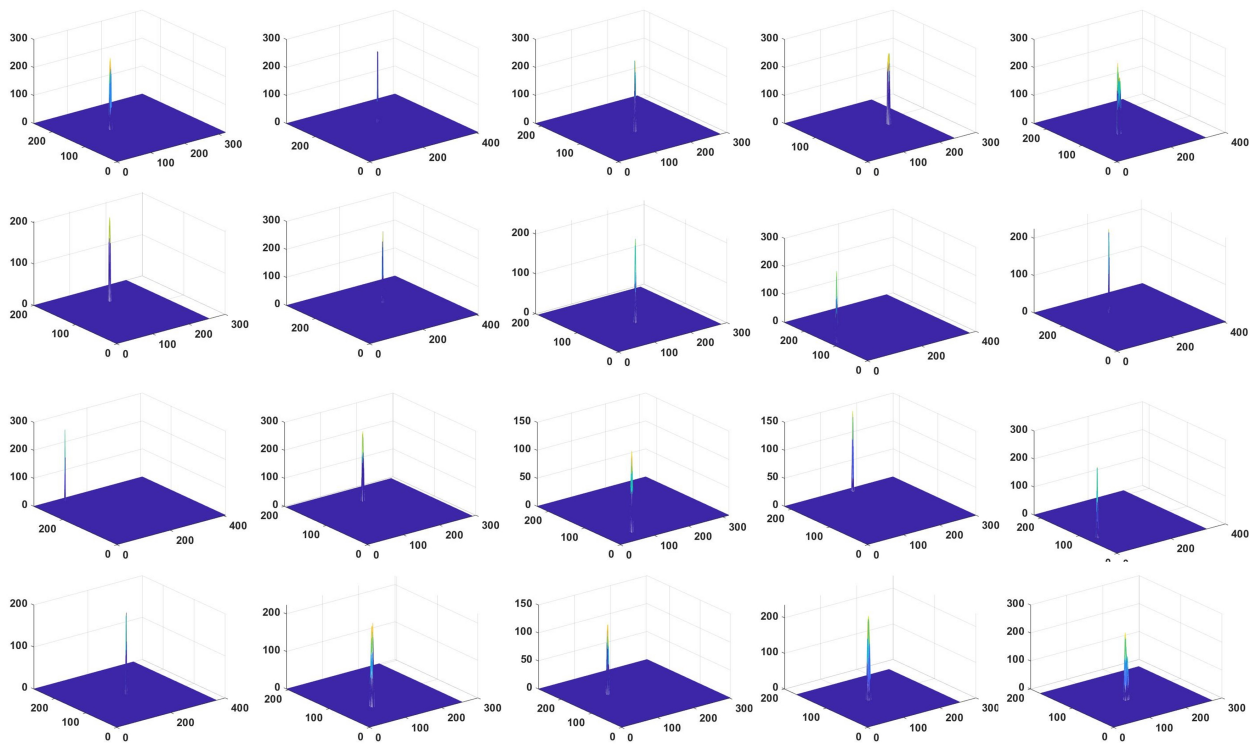


Figure 10. Three-dimensional results of the proposed method on the 20 single-frame images.

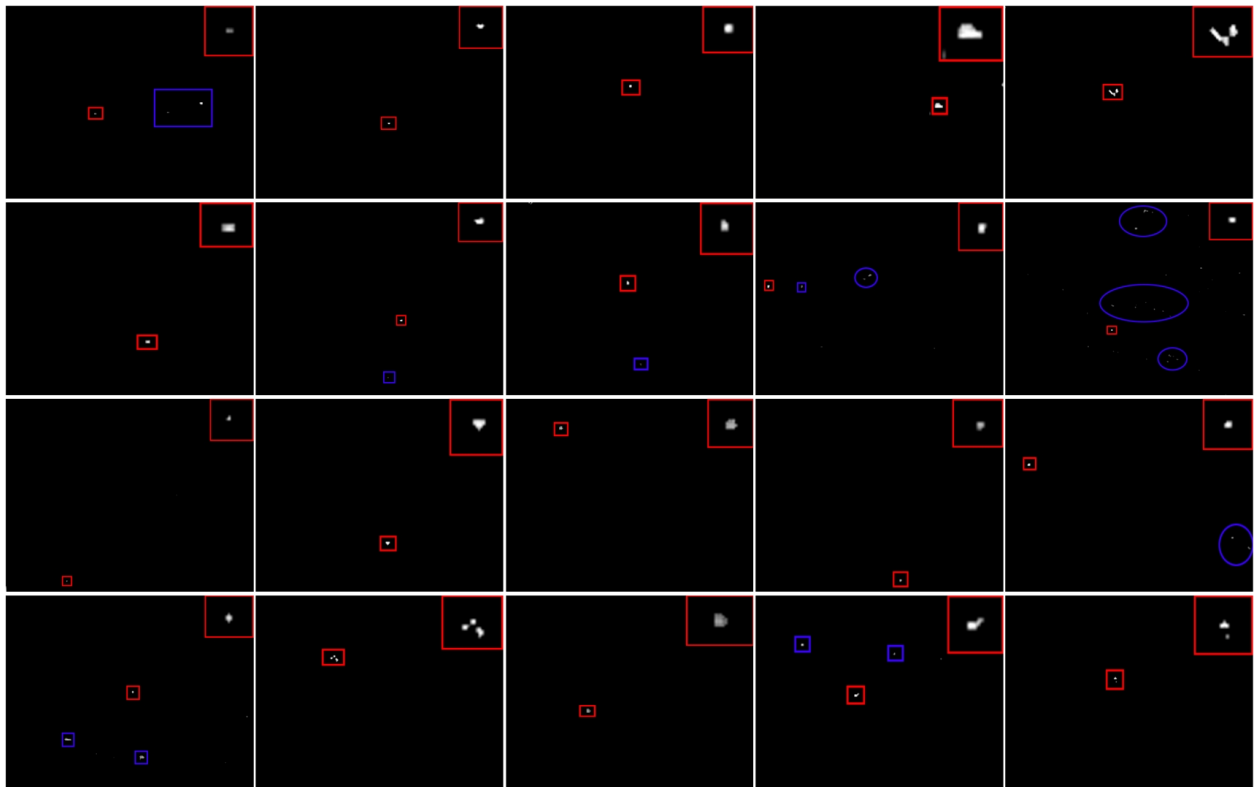


Figure 11. Twenty single-frame image results processed by PSTNN.

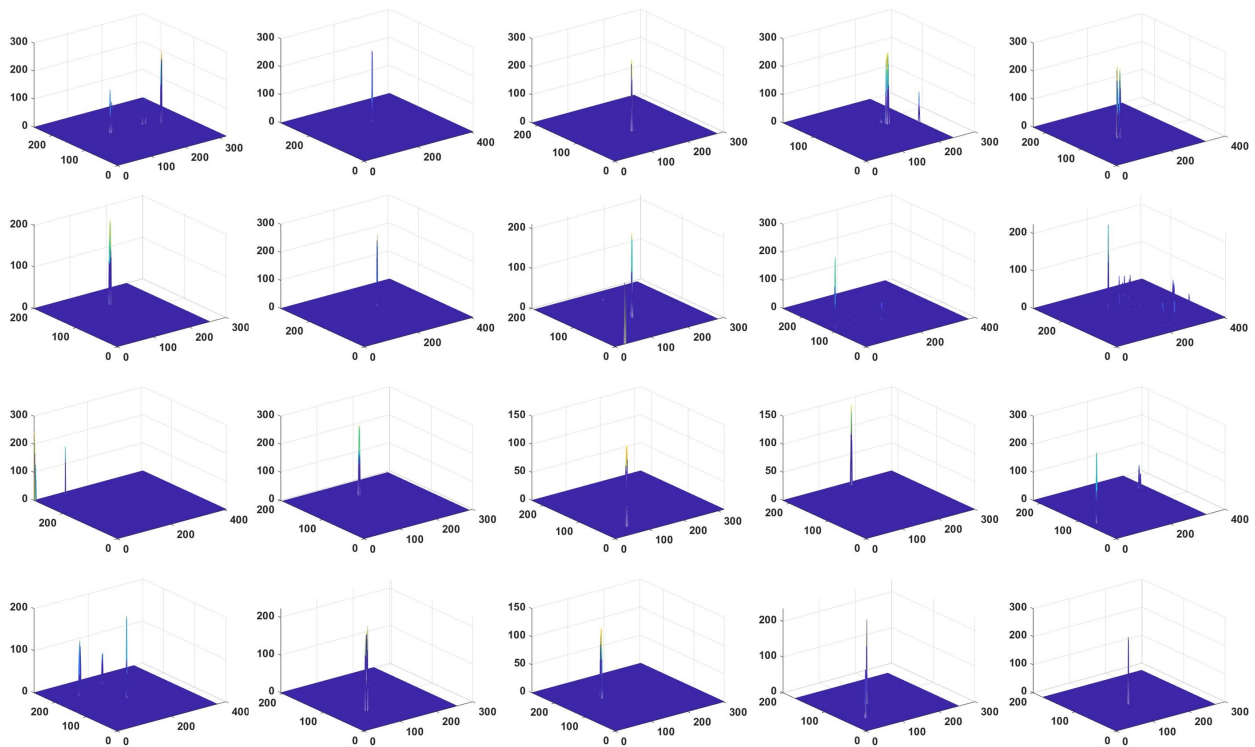


Figure 12. 3D maps of PSTNN on the 20 single-frame images.

4.5. Computation Time Comparison

In addition to assessing the detection capability of different methods, we also evaluate their computation time in this subsection. The average running time of different methods

on the six sequences are listed in Table 5. All experiments are implemented by using MATLAB R2020a on a personal computer with 2.9 GHz Ryzen7 and 16 G RAM. Among them, the Top-hat has the fastest computation speed, but its detection performance is not satisfactory, as shown in the previous experiment. Among the optimization-based algorithms, the IPI has the slowest detection speed, which is due to the fact that it employs the accelerated proximal gradient method with high time computation to solve its model. Among the IPT model-based algorithms, PSTNN provides the fastest computation speed. The proposed algorithm provides a slower computation speed than PSTNN, but it is faster than the RIPT and LogTFNN. Thus, the computing time is acceptable.

Table 5. The computing time of different methods for six sequences (Seconds).

	Top-Hat	WLCM	IPI	SMSL	RIPT	LogTFNN	PSTNN	Proposed
Seq.1	0.0041	1.9749	4.1769	0.1681	0.7069	1.2196	0.2093	0.4934
Seq.2	0.0038	2.0203	4.0588	0.1792	0.6564	1.2241	0.2108	0.4341
Seq.3	0.0026	2.0071	4.0922	0.2792	0.6307	1.2152	0.1921	0.3925
Seq.4	0.0039	1.9492	4.2002	0.3444	0.5800	1.2780	0.1804	0.4050
Seq.5	0.0037	1.9568	4.5251	0.2464	0.5607	1.2455	0.2037	0.3734
Seq.6	0.0035	1.9627	4.0498	0.2641	0.5534	1.3670	0.1853	0.5612

5. Conclusions

In this paper, a tensor optimization patch model that combines gradient features with edge and highlighted areas as prior regulations is proposed for infrared small-target detection. To accurately separate the target and background, we utilize the Laplace function to approximate the background rank. Moreover, gradient features and edge and highlighted area indicators are integrated as local targets prior to suppressing edge clutter. Then, ADMM is employed to solve the proposed model. The extensive experimental results on numerous real scene sequences demonstrate that the proposed method outperforms other baseline methods on both visual perception and indices quantitative calculations.

Author Contributions: Conceptualization, Q.S. and K.B.; methodology, Z.L. (Ziling Lu) and Q.S.; software, Z.L. (Ziling Lu); validation, Z.L. (Ziling Lu) and Z.L. (Zhengtao Li); formal analysis, K.B. and Z.L. (Zhengtao Li); investigation, K.B. and Z.H.; resources, Z.L. (Ziling Lu); data curation, Q.S. and Z.H.; writing—original draft preparation, Z.L. (Ziling Lu); writing—review and editing, Q.S., Z.H. and Z.L. (Zhengtao Li); visualization, Z.H. and Z.L. (Ziling Lu); supervision, Q.S. and K.B.; project administration, Q.S.; funding acquisition, Q.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Research Project of Education Department of Jilin Province No. JJKH20220119KJ.

Data Availability Statement: The experimental data presented in this paper are available on request from the corresponding author.

Conflicts of Interest: The authors declared that they have no conflict of interest to this work.

Appendix A

Theorem A1. *Tensor singular value decomposition (t-SVD) [68]. Given a three-dimensional tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, and its T-SVD is formulated as*

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \quad (\text{A1})$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensors and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a diagonal tensor.

Algorithm A1 shows an efficient calculation procedure of T-SVD.

Algorithm A1 T-SVD for a three-order tensor**Input:** $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ **Output:** T-SVD components \mathcal{U}, \mathcal{S} and \mathcal{V} of \mathcal{A} **Step 1.** Calculate $\overline{\mathcal{A}} = \text{fft}(\mathcal{A}, [], 3)$ **Step 2.** Calculate each frontal slice $\overline{\mathcal{U}}^{(t)}, \overline{\mathcal{S}}^{(t)}$ and $\overline{\mathcal{V}}^{(t)}$ from $\overline{\mathcal{A}}$ through

for $t = 1, 2, \dots, \lceil (n_3 + 1)/2 \rceil$ **do**
 $[\overline{\mathcal{U}}^{(t)}, \overline{\mathcal{S}}^{(t)}, \overline{\mathcal{V}}^{(t)}] = \text{SVD}(\overline{\mathcal{A}}^{(t)})$

end for**for** $t = \lceil (n_3 + 1)/2 \rceil + 1, \dots, n_3$ **do** $\overline{\mathcal{U}}^{(t)} = \text{conj}(\overline{\mathcal{U}}^{(n_3 - t + 2)})$ $\overline{\mathcal{S}}^{(t)} = \overline{\mathcal{S}}^{(n_3 - t + 2)}$ $\overline{\mathcal{V}}^{(t)} = \text{conj}(\overline{\mathcal{V}}^{(n_3 - t + 2)})$ **end for****Step 3.** Calculate $\mathcal{U} = \text{ifft}(\overline{\mathcal{U}}, [], 3)$, $\mathcal{S} = \text{ifft}(\overline{\mathcal{S}}, [], 3)$, $\mathcal{V} = \text{ifft}(\overline{\mathcal{V}}, [], 3)$ **Theorem A2.** For a proximity minimization problem in the form:

$$\arg \min_X \eta \|X\|_1 + \frac{1}{2} \|X - Y\|_F^2 \quad (\text{A2})$$

in which $\eta > 0$, and $X, Y \in \mathbb{R}^{n_1 \times n_2}$, its solution can be given by the soft thresholding operator [61], as:

$$\text{Softshrink}(y, \eta) = \text{sign}(y) \times \max(|y| - \eta, 0) \quad (\text{A3})$$

Definition A1. Tensor conjugate transpose [68]. Given a tensor $\mathcal{X} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$, then its conjugate transpose $\mathcal{X}^H \in \mathbb{C}^{n_2 \times n_1 \times n_3}$ can be obtained by the following calculation:

$$\begin{aligned} (\mathcal{X}^H)^{(1)} &= (\mathcal{X}^{(1)})^H \quad \text{and} \\ (\mathcal{X}^H)^{(t)} &= (\mathcal{X}^{(n_3 + 2 - t)})^H, \quad t = 2, \dots, n_3 \end{aligned} \quad (\text{A4})$$

References

- Liu, X.; Zuo, Z. A Dim Small Infrared Moving Target Detection Algorithm Based on Improved Three-Dimensional Directional Filtering. In Proceedings of the Chinese Conference on Image and Graphics Technologies, Beijing, China, 2–3 April 2013.
- Grossi, E.; Lops, M.; Venturino, L. A novel dynamic programming algorithm for track-before-detect in radar systems. *IEEE Trans. Signal Process.* **2013**, *61*, 2608–2619. [CrossRef]
- Li, B.; Xu, Z.; Zhang, J.; Wang, X.; Fan, X. Dim-Small Target Detection Based on Adaptive Pipeline Filtering. *Math. Probl. Eng.* **2020**, *2020*, 8234349. [CrossRef]
- Bae, T.W.; Lee, S.H.; Sohng, K.I. Small target detection using the Bilateral Filter based on Target Similarity Index. *IEICE Electron. Express* **2010**, *7*, 589–595. [CrossRef]
- Deshpande, S.D.; Meng, H.E.; Ronda, V.; Chan, P. Max-Mean and Max-Median Filters for Detection of Small-Targets. In Proceedings of the SPIE's International Symposium on Optical Science, Engineering and Instrumentation, Denver, CO, USA, 18–23 July 1999.
- Bae, T.W.; Kim, B.I.; Kim, Y.C.; Ahn, S.H. Jamming effect analysis of infrared reticle seeker for directed infrared countermeasures. *Infrared Phys. Technol.* **2012**, *55*, 431–441. [CrossRef]
- Tom, V.T.; Peli, T.; Leung, M.; Bondaryk, J.E. Morphology-based algorithm for point target detection in infrared backgrounds. In Proceedings of the Signal and Data Processing of Small Targets 1993, Orlando, FL, USA, 12–14 April 1993; pp. 2–11.
- Bai, X.; Zhou, F. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognit.* **2010**, *43*, 2145–2156. [CrossRef]
- Shao, Z.; Zhu, X.; Liu, J. Morphology infrared image target detection algorithm optimized by genetic theory. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 1299–1304.
- Song, Q.; Wang, Y.; Dai, K.; Bai, K. Single frame infrared image small target detection via patch similarity propagation based background estimation. *Infrared Phys. Technol.* **2020**, *106*, 103197. [CrossRef]

11. Huang, S.; Liu, Y.; He, Y.; Zhang, T.; Peng, A.Z. Structure-Adaptive Clutter Suppression for Infrared Small Target Detection: Chain-Growth Filtering. *Remote Sens.* **2019**, *12*, 47. [[CrossRef](#)]
12. Chen, C.; 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](#)]
13. 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](#)]
14. Hong, Z.; Lei, Z.; Ding, Y.; Hao, C. Infrared small target detection based on local intensity and gradient properties. *Infrared Phys. Technol.* **2017**, *89*, 88–96.
15. 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. [[CrossRef](#)]
16. Han, J.; Moradi, S.; Faramarzi, I.; Liu, C.; 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. [[CrossRef](#)]
17. Nasiri, M.; Chehresa, S. Infrared small target enhancement based on variance difference. *Infrared Phys. Technol.* **2017**, *82*, 107–119. [[CrossRef](#)]
18. Bai, X.; Bi, Y. Derivative Entropy-Based Contrast Measure for Infrared Small-Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2452–2466. [[CrossRef](#)]
19. Gao, C.; Meng, D.; Yang, 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](#)] [[PubMed](#)]
20. Xie, Y.; Gu, S.; Liu, Y.; Zuo, W.; Zhang, W.; Zhang, L.; Yuan, X. Weighted Schatten p -Norm Minimization for Image Denoising and Background Subtraction. *IEEE Trans. Image Process.* **2016**, *25*, 4842–4857. [[CrossRef](#)]
21. Dai, Y.; Wu, Y.; Song, Y. Infrared small target and background separation via column-wise weighted robust principal component analysis. *Infrared Phys. Technol.* **2016**, *77*, 421–430. [[CrossRef](#)]
22. 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](#)]
23. Zhang, T.; Wu, H.; Liu, Y.; Peng, L.; Yang, C.; Peng, Z. Infrared Small Target Detection Based on Non-Convex Optimization with L_p -Norm Constraint. *Remote Sens.* **2019**, *11*, 559. [[CrossRef](#)]
24. Wang, X.; Peng, Z.; Kong, D.; Zhang, P.; He, Y. Infrared dim target detection based on total variation regularization and principal component pursuit. *Image Vis. Comput.* **2017**, *63*, 1–9. [[CrossRef](#)]
25. Wang, X.; Peng, Z.; Kong, D.; He, Y. Infrared Dim and Small Target Detection Based on Stable Multisubspace Learning in Heterogeneous Scene. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5481–5493. [[CrossRef](#)]
26. Zhang, T.; Peng, Z.; Wu, H.; He, Y.; Li, C.; Yang, C. Infrared small target detection via self-regularized weighted sparse model. *Neurocomputing* **2021**, *420*, 124–148. [[CrossRef](#)]
27. Yimian, D.; Yiquan, W. Reweighted Infrared Patch-Tensor Model with Both Nonlocal and Local Priors for Single-Frame Small Target Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3752–3767.
28. Zhang, L.; Peng, Z. Infrared Small Target Detection Based on Partial Sum of the Tensor Nuclear Norm. *Remote Sens.* **2019**, *11*, 382. [[CrossRef](#)]
29. 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](#)]
30. Wang, P.; Wang, L.; Leung, H.; Zhang, G. Super-resolution mapping based on spatial–spectral correlation for spectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 2256–2268. [[CrossRef](#)]
31. Shang, X.; Song, M.; Wang, Y.; Yu, C.; Yu, H.; Li, F.; Chang, C.I. Target-constrained interference-minimized band selection for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 6044–6064. [[CrossRef](#)]
32. Chen, X.; Xie, C.; Tan, M.; Zhang, L.; Hsieh, C.J.; Gong, B. Robust and Accurate Object Detection via Adversarial Learning. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 16617–16626.
33. Dai, X.; Jiang, Z.; Wu, Z.; Bao, Y.; Wang, Z.; Liu, S.; Zhou, E. General Instance Distillation for Object Detection. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 7838–7847.
34. Oveis, A.H.; Giusti, E.; Ghio, S.; Martorella, M. A Survey on the Applications of Convolutional Neural Networks for Synthetic Aperture Radar: Recent Advances. *IEEE Aerosp. Electron. Syst. Mag.* **2022**, *37*, 18–42. [[CrossRef](#)]
35. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
36. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
37. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

38. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In *Proceedings of the Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 11–14 October 2016*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 21–37.
39. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*; pp. 2117–2125.
40. Wang, K.; Du, S.; Liu, C.; Cao, Z. Interior Attention-Aware Network for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5002013. [[CrossRef](#)]
41. Li, B.; Xiao, C.; Wang, L.; Wang, Y.; Lin, Z.; Li, M.; An, W.; Guo, Y. Dense nested attention network for infrared small target detection. *arXiv* **2022**, arXiv:2106.00487.
42. Shi, M.; Wang, H. Infrared dim and small target detection based on denoising autoencoder network. *Mob. Netw. Appl.* **2020**, *25*, 1469–1483. [[CrossRef](#)]
43. 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](#)]
44. Xu, Y.; Hao, R.; Yin, W.; Su, Z. Parallel matrix factorization for low-rank tensor completion. *Inverse Probl. Imaging* **2017**, *9*, 601–624. [[CrossRef](#)]
45. 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 (CVPR), Columbus, OH, USA, 23–28 June 2014*.
46. Sun, W.W.; Lu, J.; Liu, H.; Cheng, G. Provable sparse tensor decomposition. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2017**, *79*, 899–916. [[CrossRef](#)]
47. Chen, Y.; Song, B.; Wang, D.; Guo, L. An effective infrared small target detection method based on the human visual attention. *Infrared Phys. Technol.* **2018**, *95*, 128–135. [[CrossRef](#)]
48. Peng, D.; Hamdulla, A. Infrared Small Target Detection Using Homogeneity-Weighted Local Contrast Measure. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 514–518.
49. Deng, H.; Sun, X.; Liu, M.; Ye, C. Infrared small-target detection using multiscale gray difference weighted image entropy. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 60–72. [[CrossRef](#)]
50. Deng, H.; Sun, X.; Liu, M.; Ye, C.; Zhou, X. Small Infrared Target Detection Based on Weighted Local Difference Measure. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4204–4214. [[CrossRef](#)]
51. Gu, Y.; Wang, C.; Liu, B.X.; 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](#)]
52. 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](#)] [[PubMed](#)]
53. Pang, D.; Shan, T.; Li, W.; Ma, P.; Tao, R. Infrared Dim and Small Target Detection Based on Greedy Bilateral Factorization in Image Sequences. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3394–3408. [[CrossRef](#)]
54. Xu, W.H.; Zhao, X.L.; Ji, T.Y.; Miao, J.Q.; Ma, T.H.; Wang, S.; Huang, T.Z. Laplace function based nonconvex surrogate for low-rank tensor completion. *Signal Process. Image Commun.* **2019**, *73*, 62–69. [[CrossRef](#)]
55. Zhang, P.; Zhang, L.; Wang, X.; Shen, F.; 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](#)]
56. Li, Z.; Zheng, J.; Zhu, Z.; Yao, W.; Wu, S. Weighted guided image filtering. *IEEE Trans. Image Process.* **2015**, *24*, 120–129.
57. Fang, H.; Chen, M.; Liu, X.; Yao, S. Infrared Small Target Detection with Total Variation and Reweighted Regularization. *Math. Probl. Eng.* **2020**, *2020*, 1529704. [[CrossRef](#)]
58. Zhou, F.; Wu, Y.; Dai, Y.; Wang, P. Detection of small target using Schatten 1/2 quasi-norm regularization with reweighted sparse enhancement in complex infrared scenes. *Remote Sens.* **2019**, *11*, 2058. [[CrossRef](#)]
59. Gu, S.; Xie, Q.; Meng, D.; Zuo, W.; Feng, X.; Zhang, L. Weighted Nuclear Norm Minimization and Its Applications to Low Level Vision. *Int. J. Comput. Vis.* **2017**, *121*, 183–208. [[CrossRef](#)]
60. Wang, H.; Yang, F.; Zhang, C.; Ren, M. Infrared small target detection based on patch image model with local and global analysis. *Int. J. Image Graph.* **2018**, *18*, 1850002. [[CrossRef](#)]
61. Hale, E.T.; Yin, W.; Zhang, Y. Fixed-point continuation for ℓ_1 -minimization: Methodology and convergence. *SIAM J. Optim.* **2008**, *19*, 1107–1130. [[CrossRef](#)]
62. Wang, Q.; Zhang, L.; Wei, S.; Li, B.; Xi, Y. Tensor Factorization-based Particle Swarm Optimization for Large-Scale Many-Objective Problems. *Swarm Evol. Comput.* **2021**, *69*, 100995. [[CrossRef](#)]
63. Chen, Y.; Guo, Y.; Wang, Y.; Dong, W.; Chong, P.; He, G. Denoising of Hyperspectral Images Using Nonconvex Low Rank Matrix Approximation. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5366–5380. [[CrossRef](#)]
64. Jie, L.; He, Z.; Chen, Z.; Lei, S. Tiny and Dim Infrared Target Detection Based on Weighted Local Contrast. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1780–1784.
65. Hui, B.; Song, Z.; Fan, H.; Zhong, P.; Hu, W.; Zhang, X.; Ling, J.; Su, H.; Jin, W.; Zhang, Y.; et al. A dataset for infrared detection and tracking of dim-small aircraft targets under ground/air background. *China Sci. Data* **2020**, *5*, 291–302.

-
66. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Asymmetric Contextual Modulation for Infrared Small Target Detection. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2021; pp. 949–958.
 67. Wang, H.; Zhou, L.; Wang, L. Miss Detection vs. False Alarm: Adversarial Learning for Small Object Segmentation in Infrared Images. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019; pp. 8508–8517.
 68. Kilmer, M.E.; Martin, C.D. Factorization strategies for third-order tensors. *Linear Algebra Appl.* **2011**, *435*, 641–658. [[CrossRef](#)]