Infrared Dim Small Target Detection via Fractional-order TV and Nonnegative L1-norm

Yebin Chen, Guimeng Cao, and Wei Xue, *Member, IAENG*

***Abstract*—Detection of dim infrared small targets is essential for defense missions, including reconnaissance, surveillance, long-range precision engagement, and airborne protection. However, distinguishing these targets against cluttered, low-contrast backgrounds remains a formidable challenge. The thesis presents a novel and efficient model specifically designed for detecting dim infrared small targets in complex environments. First, the input infrared image is decomposed into three patch-based matrices corresponding to background, target, and noise components. In order to weaken the edge information in the background, the thesis applies a nonconvex low-rank approximation to the background block matrix for suppression. Since complex backgrounds may produce pronounced residual edges in the estimated target component, fractional total variation (FTV) regularization is incorporated to capture background characteristics more comprehensively, thereby enhancing background estimation. Furthermore, a non-negative L1 norm constraint is imposed on the target matrix to preserve weak and dim signals. Subsequently, the resulting minimization problem is addressed using an alternating optimization strategy. Comprehensive evaluations on a real-world low-altitude drone dataset indicate that our approach outperforms existing solutions, yielding enhanced visual and numerical results for detecting faint infrared small targets.**

***Index Terms*—Dim infrared small targets, Nonconvex low-rank approximation, Fractional total variation, Non-negative L1 norm.**

I. Introduction

Infrared dim small target detection, a vital subfield of target recognition, finds extensive applications in precision guidance, spaceborne surveillance, unmanned aerial vehicle detection, and sensitive target tracking. It has significant practical value, particularly in the military, security, and aerospace sectors. Infrared imaging detects the thermal radiation emitted by objects, independent of lighting conditions, enabling effective monitoring both day and night, in adverse weather conditions, and in complex environments. However, these targets have weak radiation energy and suffer from low signal-to-noise ratio (SNR) and signal-to-clutter ratio (SCR), making them susceptible to being overwhelmed by noise in complex backgrounds [1]. In addition, long-range imaging results in the loss of texture and contour details of the target, which typically appears as small, blurry spots lacking distinct edges in infrared images, thereby increasing the difficulty of detection. In recent years, with ongoing advancements in infrared sensor technology, the demand for the detection of weak and small infrared targets has become increasingly urgent, leading to the emergence of various detection algorithms. Single-frame detection algorithms are commonly classified into four categories: background consistency methods, target saliency methods, deep-learning–based methods, and mathematical-optimization–based methods.

Background consistency methods assume that background transitions are slow and local background similarity is high. These methods focus on suppressing gentle background clutter and indirectly enhancing the target. Common algorithms include Tophat [2], Max-Mean, and Max-Median [3], among others. These algorithms suppress background clutter by designing filter kernels. A drawback of this approach is that the fixed filter template lacks adaptability to complex scenarios, yielding satisfactory performance only under simple background conditions. As a result, background residues frequently remain in the detection outputs, leading to an elevated false alarm rate. Target saliency methods are based on the saliency of objects, inspired by the human visual system. The human eye tends to focus on salient targets, and weak targets are often abrupt points on a gentle background in infrared images. Therefore, some researchers have introduced contrast mechanisms, adaptive scaling, and other aspects of the human visual system into detection algorithms. LoG (Laplacian of Gaussian) [4], LCM (local contrast method) [5], and others are classic saliency algorithms. In recent years, more complex measurement methods have been proposed, such as derivative difference measurement [6], dual-layer local contrast measurement [7], and spatio-temporal local difference measurement [8]. Such algorithms require the target to exhibit certain characteristics. However, real-world scenes are complex, and these algorithms perform poorly when high radiation interference sources are present. Learning-based approaches leverage model training to automatically distinguish feature differences between targets and backgrounds in infrared imagery, emerging as a key research direction in infrared small target detection. These approaches are broadly categorized into traditional machine learning techniques and deep learning methods. Traditional methods rely on feature extractors (e.g., gradients, textures, local contrast) and classifiers (e.g., SVM and random forests), offering low computational complexity but poor adaptability to complex backgrounds. Deep learning methods automatically extract multi-scale features through end-to-end training, with stronger background suppression and target enhancement capabilities. HCF-Net (Hierarchical Context Fusion Network) [9] improves detection performance

Manuscript received December 26, 2024; revised June 6, 2025. This work was supported in part by the Natural Science Foundation of the Anhui Higher Education Institutions of China (2022AH050310, 2023AH040149) and the Anhui Provincial Natural Science Foundation (2208085MF168).

Yebin Chen is a Professor of the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (e-mail: cyb7102@163.com).

Guimeng Cao is a Postgraduate Student at the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (e-mail: cguimeng@163.com).

Wei Xue is an Associate Professor of the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (Corresponding author; e-mail: xuewei@ahut.edu.cn).

through parallel attention and channel refinement modules. PBT (Progressive Background-Aware Transformer) [10] uses an asymmetric encoder-decoder structure and performs well in complex scenes. The Improved Dense Nested Attention Network (IDNANet) [11] incorporates a hierarchical vision transformer backbone with a weighted loss strategy to mitigate foreground-background class imbalance. You Only Look Once version 7X Plus (YOLOv7X+) [12] enhances the YOLOv7 structure to improve small target detection, especially in aerial images. Optimization approaches for dim infrared small-target detection often leverage intrinsic image characteristics by representing the background as a low-rank component and the target as a sparse element. By leveraging these assumptions, optimization models with suitable constraints are iteratively solved, enabling reliable target isolation from complex backgrounds. In recent years, increasing attention has been given to leveraging structural priors in infrared images, especially approaches grounded in robust principal component analysis (RPCA) [13], [14]. RPCA presumes that the background is represented by a low-rank matrix, whereas the target manifests as a sparse component within the image. Gao et al. [15] extended this idea by introducing the infrared patch-image (IPI) model, which casts detection as a decomposition of the data into low-rank background and sparse target components. However, in complex scenes, the application of nuclear-norm and L_1-norm regularizations in the IPI framework can cause loss of target details or leave background artifacts—such as prominent structural edges—within the estimated target component. To mitigate this, the weighted IPI (WIPI) model [16] proposed assigning adaptive weights to each column of the target patch matrix based on local structural properties. However, the weight estimation in WIPI is computationally expensive and time-consuming. To improve efficiency, the reweighted IPI (RIPI) model [17] introduced weighted nuclear norm and weighted L_1-norm regularizations to better adapt to varying background and target characteristics. Still, as the nuclear norm effectively acts as an L_1 penalty on the singular values, it may yield biased estimates. To address this limitation, non-convex surrogates [18] and the Schatten-$\frac{1}{2}$ norm [19] were explored for more accurate low-rank background modeling [20]. In parallel, total variation (TV) regularization [21], [22], [23] has demonstrated effectiveness in preserving both smooth and edge structures in background estimation. Based on this, the TVIPI model [24] integrated TV regularization into the IPI framework to enhance detection performance and robustness. However, its performance can degrade in cases with insufficient edge samples, leading to residual background sparsity [25]. To address this, the reweighted TVIPI (ReTVIPI) model [25] introduced a combination of spatial and temporal features. Further developments extended these models to the tensor domain, such as the tensor-based ReTVIPI framework [26], which was found effective in handling more complex background textures. Additionally, a dual-window local contrast enhancement method followed by multiscale IPI feature extraction was proposed in [27] to improve contrast between targets and backgrounds. Beyond IPI-based techniques, various other strategies have been developed to identify weak small targets in infrared images [6], [28], [29], [30], [31]. These approaches represent diverse modeling strategies and collectively advance the development of infrared small target detection technologies.

A literature survey reveals that existing detection models predominantly use the nuclear norm of a matrix as an approximation of its rank and the L_1- norm to approximate the L_0-norm during optimization. These simplistic approximations can lead to detection errors, particularly when strong edges in the background are erroneously classified into the sparse component, increasing the false alarm rate. To address these issues, the thesis proposes a new detection model based on fractional-order total variational regularization and nonnegative L_1-norm. With appropriate algorithmic optimization, the model can automatically separate the target from the background. The contributions of this work are as follows:

1) To eliminate suppress the background edge features, the thesis adopts a low-rank approximation of the background fast matrix using a non-convex γ norm
2) To eliminate the prominent residual edges in the target image under challenging backgrounds, this study combines fractional total variation regularization to more comprehensively capture background features.
3) Recognizing that true infrared small targets exhibit consistently positive intensity relative to their surroundings, we impose a nonnegativity constraint on the target patch matrix to improve detection accuracy.

II. PRELIMINARIES

A. IPI Model

The IPI model [15] performs low-rank sparse optimization decomposition by rearranging a single frame image into an infrared block image as the input image. The core idea is to decompose the infrared block image I into three block matrices B (background), T (target), and N (noise) expressed as:

$$I = B + T + N \tag{1}$$

Following the RPCA framework, the rank of B is relaxed to its nuclear norm $\|B\|_*$, and the L_0-norm of T is replaced by the L_1-norm $\|T\|_1$. We further assume that the noise matrix N has independent, identically distributed entries and satisfies

$$\|N\|_F \leq \delta, \quad \delta > 0.$$

Equivalently,

$$\|I - B - T\|_F \leq \delta.$$

Accordingly, the IPI optimization problem can thus be reformulated as:

$$\begin{aligned} &\min_{B,T} \ \|B\|_* \ + \ \lambda \|T\|_1 \\ &\text{s.t.} \quad \|I - B - T\|_F \leq \delta, \end{aligned} \tag{2}$$

Where, $\|\cdot\|_*$, $\|\cdot\|_1$, and $\|\cdot\|_F$ represent the nuclear norm, L_1-norm, and Frobenius norm, respectively.

B. γ norm

The IPI model may retain residual structural edges because it relies on the nuclear norm for rank approximation. To address this limitation, a more accurate surrogate—known

as the gamma norm—is employed to approximate the rank of matrix B [18], expressed as:

$$\|B\|_\gamma = \sum_i \frac{(1+\gamma)\sigma_i(B)}{\gamma + \sigma_i(B)}, \quad \gamma > 0 \tag{3}$$

where $\sigma_i(B)$ is the singular value of matrix B, and $\lim_{\gamma\to\infty}\|B\|_\gamma = \|B\|_*$ and $\lim_{\gamma\to 0}\|B\|_\gamma =$ rank(B) can be derived using the limit in mathematics. As $\gamma \to 0$, the γ-norm converges to $\mathrm{rank}(B)$, whereas as $\gamma \to \infty$, it converges to $\|B\|_*$. Consequently, the γ-norm interpolates between the matrix rank and its nuclear norm, providing a more accurate representation of the low-rank nature of the background matrix than the nuclear norm. Directly minimizing the rank of a matrix leads to a non-convex and computationally intractable problem. To address this, the proposed method employs the γ-norm as a tractable surrogate for rank approximation.

C. Target Nonnegativity

In prior works [15], [16], the L_1 norm characterizes the target patch-image representation. This regularization, however, exclusively captures target sparsity while disregarding the inherent physical property that infrared small targets demonstrate higher intensity than their local surroundings. Consequently, authentic targets consistently exhibit positive values within the patch-image domain. A more physically consistent approach therefore necessitates imposing a non-negativity constraint on the target patch-image, formally expressed as:

$$\|T\|_{1,\geq 0} = \sum_{i,j} \max(T_{i,j}, 0). \tag{4}$$

III. Proposed Method

This section introduces a novel IPI-based detection framework designed to enhance detection precision. The methodology, delineated in Fig. 1, comprises two primary stages: patch-image construction and image reconstruction. The reconstruction phase is subsequently partitioned into background restoration and target recovery components. Target image reconstruction is achieved through optimization of the proposed model detailed below.

A. Model

The fractional order total variation (FTV) model [32], [33] extends the classical total variation (TV) model by incorporating fractional derivatives to better handle image edges and textures while reducing noise. Classical TV focuses on first-order gradients, while FTV introduces fractional derivatives, which capture non-local image information. This enhances the capability of preserving image features such as edges and fine details that may be smoothed out in traditional TV models.

$$FTV(X) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1}\sqrt{(D_x^\alpha x_{i,j})^2 + (D_y^\alpha x_{i,j})^2} \tag{5}$$

Let $D_x^\alpha x_{i,j}$ and $D_y^\alpha x_{i,j}$ denote the fractional derivatives of $x_{i,j}$ in the horizontal and vertical directions, respectively, defined as follows:

$$\begin{aligned}(D_x^\alpha u)_{i,j} &= \sum_{k=0}^{K-1}(-1)^k C_k^\alpha u_{i-k,j}\\ (D_y^\alpha u)_{i,j} &= \sum_{k=0}^{K-1}(-1)^k C_k^\alpha u_{i,j-k}\end{aligned} \tag{6}$$

where K indicates the neighborhood size, and the fractional coefficients C_k^α are defined by:

$$C_k^\alpha = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha+1-k)} \tag{7}$$

The Gamma function $\Gamma(x)$, extending factorials to non-integer domains, is expressed as:

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt \tag{8}$$

Defining $D_i^\alpha X \in \mathbb{R}^2$ as the fractional discrete gradient of vectorized image X at pixel i, the fractional total variation (FTV) norm becomes:

$$\mathrm{FTV}(X) = \sum_i \|D_i^\alpha X\|_2, \tag{9}$$

Minimizing the FTV norm effectively denoises smooth regions while preserving edge sharpness. Unlike classical TV, FTV regulates smoothness via the fractional order α. This yields the nonnegative fractional-order TV regularized infrared patch-image model:

$$\begin{aligned}&\min_{T,B}\|B\|_\gamma + \lambda_1\mathrm{FTV(B)} + \lambda_2\|\boldsymbol{T}\|_{1,\geq 0}\\ &\text{s.t.} I = T + B + N, \|N\|_F \leq \delta.\end{aligned} \tag{10}$$

Here, $FTV(\cdot)$ represents the FTV norm, N denotes noise, and λ_1, λ_2 balance the regularization terms. Substituting Eq. (9) for FTV(B) produces the unified formulation:

$$\begin{aligned}&\min_{T,B}\|B\|_\gamma + \lambda_1\sum_i\|D_i^\alpha B\|_2 + \lambda_2\|\boldsymbol{T}\|_{1,\geq 0}\\ &\text{s.t.} I = T + B + N, \|N\|_F \leq \delta.\end{aligned} \tag{11}$$

Here, D_i^α denotes the fractional gradient operator. Applying both the Gamma norm and the FTV norm to the background patch B attenuates clutter and improves the accuracy of background estimation. Solving Eq. (11) yields the infrared target image, with subsequent postprocessing extracting target information. The complete methodology is illustrated in Fig. 1.

B. Optimization

To efficiently solve the optimization model, we develop an algorithm based on variable splitting and subproblem decomposition. The solution strategy begins by decoupling correlated variables in Eq. (11), reformulating it as an equivalent constrained optimization problem:

$$\begin{aligned}&\min_{Z_1,Z_2,Z_3}\|Z_1\|_\gamma + \lambda_1\sum_i\|z_i\|_2 + \lambda_2\|Z_3\|_{1,\geq 0}\\ &\text{s.t.}\quad Z_1 = B\\ &\qquad Z_2 = [z_1; z_2; ...; z_{mn}], \quad z_i = D_i^\alpha B\\ &\qquad Z_3 = T\\ &\qquad I = T + B + N, \quad \|N\|_F \leq \delta.\end{aligned} \tag{12}$$

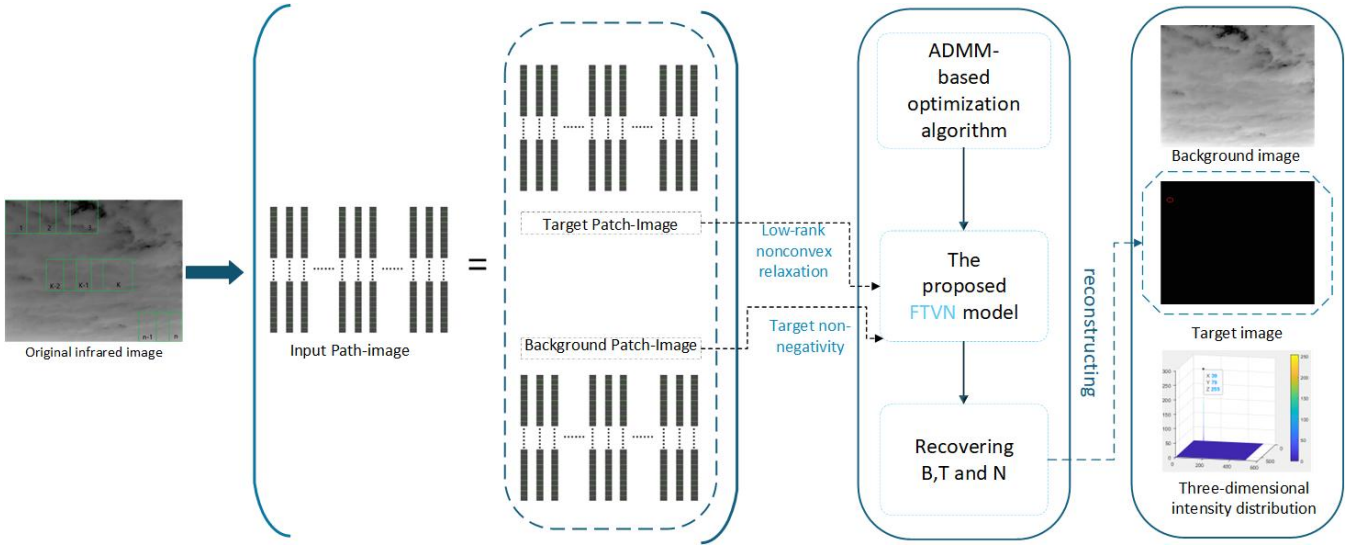

Fig. 1: An overview of the FTVN model proposed in the thesis is provided.

The corresponding augmented Lagrangian formulation for Eq. (12) is given by:

$$\begin{aligned} L_A = &\|Z_1\|_\gamma + \lambda_1 \sum_i \|z_i\|_2 + \lambda_2 \|Z_3\|_{1,\geq 0} + \langle Y_1, Z_1 - B\rangle \\ &+ \langle Y_3, Z_3 - T\rangle + \frac{\beta}{2}\|Z_3 - T\|_F^2 + \frac{\beta}{2}\|Z_1 - B\|_F^2 \\ &+ \sum_i \left(\langle y_i, z_i - D_i^\alpha B\rangle + \frac{\beta_i}{2}\|z_i - D_i^\alpha B\|_F^2 \right) \\ &+ \langle Y_4, I - T - B - N\rangle + \frac{\beta}{2}\|I - T - B - N\|_F^2. \end{aligned} \tag{13}$$

The notation $\langle \cdot, \cdot \rangle$ refers to the inner product between matrices, and Y_i $(i = 1, 2, 3, 4)$ are the Lagrange multipliers, where Y_i $(i = 1, 3, 4) \in \mathbb{R}^{mn\times 1}$ and $Y_2 = [y_1, y_2, \ldots, y_{mn}] \in \mathbb{R}^{2\times mn}$. The scalar β serves as the penalty parameter influencing the convergence behavior. To enhance computational efficiency, the variables Z_1, Z_3, T, B, and N in Eq. (13) are converted into column-vector form. The optimization is performed using the ADMM [34], [35], which iteratively updates the subproblems corresponding to Z_i $(i = 1, 2, 3)$, followed by sequential optimization of N, B, and T.After each iteration, the Lagrange multipliers undergo an update, while other variables are held constant in their respective steps.

The optimization subproblem corresponding to Z_1 is reformulated as:

$$\begin{aligned} Z_1^{k+1} &= \arg\min_{Z_1} L_A\left(Z_1, Z_2^k, Z_3^k, B^k, T^k, N^k\right) \\ &= \arg\min_{Z_1} \|Z_1\|_\gamma + \left\langle Y_1^k, Z_1 - B^k\right\rangle + \frac{\beta}{2}\left\|Z_1 - B^k\right\|_F^2 \\ &= \arg\min_{Z_1} \|Z_1\|_\gamma + \frac{\beta}{2}\left\|Z_1 - \left(B^k - \frac{Y_1^k}{\beta}\right)\right\|_F^2. \end{aligned} \tag{14}$$

To facilitate solving (14), we utilize the following lemma established by Kang et al. [18]:

Consider a matrix $A \in \mathbb{R}^{m\times n}$ with singular value decomposition given by $A = U\Sigma_A V'$, where $\Sigma_A = \mathrm{diag}(\sigma_A)$ and $F(Z)$ denotes a unitarily invariant function defined as the composition $f \circ \sigma(Z)$. The subsequent optimization problem is formulated as:

$$\min_{Z\in\mathbb{R}^{m\times n}} F(Z) + \frac{\mu}{2}\|Z - A\|_F^2 \tag{15}$$

with $\mu > 0$ admits optimal solution $Z_* = U\Sigma_Z^* V'$ where $\Sigma_Z = \mathrm{diag}(\mathrm{prox}_{f,\mu}(\sigma_A))$. The Moreau-Yosida operator is defined as

$$\mathrm{prox}_{f,\mu}(\sigma_A) \triangleq \arg\min_{\sigma\geq 0} f(\sigma) + \frac{\mu}{2}\|\sigma - \sigma_A\|_2^2. \tag{16}$$

In our model, incorporating the γ-norm renders the objective in (17) a mixture of convex and concave components. This thesis utilizes the DC (difference-of-convex) programming technique [36] to solve the problem by representing the objective function as the difference of two convex functions and linearizing the non-convex part using a first-order Taylor expansion. During the $(t+1)$-th inner iteration, the algorithm solves the following subproblem:

$$\sigma^{t+1} = \arg\min_{\sigma\geq 0} \langle w_t, \sigma\rangle + \frac{\mu_k}{2}\|\sigma - \sigma_A\|_2^2 \tag{17}$$

which yields a closed-form solution:

$$\sigma^{t+1} = \max\left(\sigma_A - \frac{w_t}{\mu_k}, 0\right).$$

Here, $A = B^k - Y_1^k/\beta$, and $w_t = \gamma(1+\gamma)/(\gamma + \sigma^t)^2$ represents the gradient of f at σ^t. Once convergence to a local optimum σ_* is achieved, the update $Z_1^{k+1} = U\mathrm{diag}(\sigma_*)V'$ is used to solve (14).

The optimization problem for Z_2 is equivalent to:

$$\begin{aligned} Z_2^{k+1} &= \arg\min_{Z_2} L_A\left(Z_1^k, Z_3^k, B^k, T^k, N^k\right) \\ &= \arg\min_{Z_2} \sum_i \left(\|z_i\|_2 + \left\langle y_i^k, z_i - D_i^\alpha B^k\right\rangle \right. \\ &\quad \left. + \frac{\beta_i}{2}\left\|z_i - D_i^\alpha B^k\right\|_F^2\right). \end{aligned} \tag{18}$$

It is worth noting that this is a constrained L2 optimization problem, solvable via a two-dimensional shrinkage-based

solution

$$z_i = \max\left\{\left\|D_i^\alpha B - \frac{y_i}{\beta_i}\right\|_2 - \frac{1}{\beta_i}, 0\right\} \cdot \frac{\left(D_i^\alpha B - \frac{y_i}{\beta_i}\right)}{\left\|D_i^\alpha B - \frac{y_i}{\beta_i}\right\|_2}. \tag{19}$$

The Z_3 subproblem aims to solve the following optimization formulation:

$$\begin{aligned} Z_3^{k+1} &= \arg\min_{Z_3} L_A\left(Z_1^k, Z_2^k, B^k, T^k, N^k\right) \\ &= \arg\min_{Z_3} \lambda_2 \|Z_3\|_1 + \frac{\beta}{2}\left\|Z_3 - \left(T^k - \frac{Y_3^k}{\beta}\right)\right\|_F^2 \end{aligned} \tag{20}$$

Eq. (20) admits the following closed-form solution:

$$Z_3^{k+1} = Th_{\frac{\lambda_2}{\beta}}\left(T - \frac{Y_3^k}{\beta}\right) \tag{21}$$

$$Th_\varepsilon(W) = \begin{cases} w - \varepsilon & w > \varepsilon \\ w + \varepsilon & w < -\varepsilon \\ 0 & otherwise \end{cases} \tag{22}$$

where $Th_\varepsilon(\cdot)$ is the thresholding operator [37].

In the second step, the subproblems for N, B, and T are addressed. The N-subproblem can be reformulated as follows:

$$\begin{aligned} N^{k+1} &= \arg\min_N \left\langle Y_4^k, I - T^k - B^k - N\right\rangle \\ &\quad + \frac{\beta}{2}\left\|I - T^k - B^k - N\right\|_F^2 \\ &= \left\|N - \left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right)\right\|_F^2, \\ &\text{s.t. } \|N\|_F \le \delta. \end{aligned} \tag{23}$$

The solution to Eq. (23) can be obtained using the following projection operator:

$$N^{k+1} = P_\Omega\left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right). \tag{24}$$

where Ω denotes the set of matrices with Frobenius norm bounded by δ, and P_Ω represents the orthogonal projection onto Ω. The B-subproblem can be solved by:

$$B^{k+1} \leftarrow \frac{\partial L_A}{\partial B} = 0. \tag{25}$$

Equation (25) is linear and can be expressed equivalently as:

$$\begin{aligned} B^{k+1} = &\left(2\beta + \sum_i \beta_i (D_i^\alpha)^T D_i^\alpha\right)^{-1} \\ &\cdot \left[Y_1^k + Y_4^k + \sum_i \left((D_i^\alpha)^T y_i + \beta_i (D_i^\alpha)^T z_i\right)\right. \\ &\left. + \beta\left(Z_1^{k+1} + I - T^k - N^{k+1}\right)\right] \end{aligned} \tag{26}$$

The solution to the T-subproblem is given by:

$$T^{k+1} \leftarrow \frac{\partial L_A}{\partial T} = 0 \tag{27}$$

$$T^{k+1} = \frac{Y_3 + \beta Z_3^{k+1} + Y_4 + \beta\left(I - B^{k+1} - N^{k+1}\right)}{2\beta} \tag{28}$$

During the third phase, the Lagrange multipliers are iteratively adjusted based on the updated optimization variables

$$\begin{aligned} Y_1^{k+1} &\leftarrow Y_1^k + \mu\beta\left(Z_1^{k+1} - B^{k+1}\right) \\ Y_2^{k+1} &\leftarrow Y_2^k + \mu\beta\left(Z_2^{k+1} - D^\alpha B^{k+1}\right) \\ Y_3^{k+1} &\leftarrow Y_3^k + \mu\beta\left(Z_3^{k+1} - T^{k+1}\right) \\ Y_4^{k+1} &\leftarrow Y_4^k + \mu\beta\left(I - T^{k+1} - B^{k+1} - N^{k+1}\right). \end{aligned} \tag{29}$$

where $\mu > 0$ is a predetermined step size that eliminates the need to update β at each iteration, thereby enhancing the stability of the solution. Algorithms 1 and 2 summarize the overall optimization procedure.

Algorithm 1: FTVN Model Optimization Procedure

Input: Input patch $I \in \mathbb{R}^{m\times n}$; parameters β, λ_1, λ_2, tolerance δ

Output: Estimated background B_k and target T_k

Initialize: $B = T = N = Z_1 = Z_3 = Z_4 = Y_1 = Y_3 = Y_4 = \text{zeros}(m, n)$,

vectorize $Z_2 = \text{zeros}(mn, 2)$, $Y_2 = \text{zeros}(2, mn)$.

while *stopping criteria not met* **do**

- Update Z_1^{k+1} as per Algorithm 2 and Eq. (17);
- Compute Z_2^{k+1} using the 2-D shrinkage operator (Eq. (19));
- Compute $Z_3^{k+1} = \text{Th}_{\frac{\lambda_2}{\beta}}\left(T^k - \frac{Y_3^k}{\beta}\right)$;
- Update $N^{k+1} = P_\Omega\left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right)$;
- Update B^{k+1} following Eq. (26);
- Update T^{k+1} based on Eq. (28);
- Refresh multipliers Y_i^{k+1} for $i = 1, 2, 3, 4$ using Eq. (29);
- Verify whether the convergence criterion is satisfied:

$$\frac{\|I - B^{k+1} - T^{k+1} - N^{k+1}\|_F}{\|I\|_F} < \delta$$

- Increment iteration counter: $k \leftarrow k + 1$;

Algorithm 2: Optimization of Non-Convex Problems via DC Programming

Input: Initialization parameters B_k, Δ_0, Λ, γ, ϵ_3

Output: $B_{k+1} = U\Delta_{t+1}V$

Initialization: Set $B_0 = T_0 = N_0 = X_0 = \text{zeros}(m, n)$, choose ϵ_2, and $t = 0$;

while *the convergence condition remains unmet* **do**

- Calculate weights $w_t = \frac{\gamma(1+\gamma)}{(\gamma+\Delta_t)^2}$ and update $\Delta_{t+1} = \max(0, S - \frac{w_t}{\mu})$;
- if $\|\Delta_{t+1} - \Delta_t\|_2 < \epsilon_3$, then stop;
- Increment iteration count $t \leftarrow t + 1$;

IV. Experiments

This section presents an evaluation of the proposed method's detection performance using real-world datasets.

A. Evaluation Metrics

We adopted an evaluation framework that combines both qualitative and quantitative metrics, including TDID (3D intensity distribution), BSF (background suppression factor), SNRG (signal-to-noise ratio gain), and computational efficiency. All experiments were conducted on a Windows 11 platform equipped with an Intel Core™ i5-13600KF CPU (3.50 GHz), 32 GB of RAM, and MATLAB R2024a.

The target-to-interference difference (TDID) metric characterizes an image's gray-level distribution, utilized here to visually assess the clutter suppression quality in detected target images.

The BSF measures the effectiveness of background attenuation, with larger BSF values indicating stronger suppression. It is defined as:

$$\mathrm{BSF} = \frac{C_{\mathrm{in}}}{C_{\mathrm{out}}} \tag{30}$$

Let C_{in} and C_{out} denote the standard deviations of the background region in the input and processed images, respectively. A higher BSF value indicates a more effective suppression of background fluctuation.

SNRG measures target enhancement effectiveness:

$$\mathrm{SNRG} = \frac{\mathrm{SNR}_{\mathrm{out}}}{\mathrm{SNR}_{\mathrm{in}}} \tag{31}$$

The terms $\mathrm{SNR}_{\mathrm{in}}$ and $\mathrm{SNR}_{\mathrm{out}}$ quantify the image signal-to-noise ratios before and after applying the model, respectively. The SNR is computed as $(G_{\max} - G_{\mathrm{mean}})/\theta$, where $G_{\max}$, G_{mean}, and θ correspond to the maximum gray value, mean gray value, and standard deviation, respectively. Higher SNRG values indicate better target enhancement capability.

B. Experimental Setup

Six representative datasets were chosen from infrared small-target repositories to assess the detection performance of the proposed method. Their detailed characteristics are listed in Table I.

To evaluate its effectiveness, the proposed method is compared with eight advanced techniques for dim infrared small target detection, including two traditional baselines (MAXMED and TOPHAT) and six recent algorithms (PSTNN, IPIAPG, TV-PCP, NRAM, LAMPS, and 4D-TR). The parameter configurations for these methods are summarized in Table II.

During the patch-image construction process, overlapping patches may result in multiple pixel assignments in the reconstructed image. Consistent with [15], to mitigate this, a one-dimensional median filtering technique is utilized to reconstruct the background and target components.

C. Results and Analysis

This section presents the comparative results of the nine evaluated approaches. We begin by specifying the parameter settings for the proposed FTVN method. Extensive experiments reveal that the optimal performance of the proposed model is achieved with a block size of 50×50 and a sliding stride of 15. The parameter λ_1 is chosen empirically and typically set around 0.01; in our experiments, the parameter λ_1 was fixed at 0.005. The balancing parameter λ_2, which mediates the trade-off between the γ-norm and L_1-norm contributions while incorporating the FTV term's influence, was selected as:

$$\lambda_2 = \frac{1}{\sqrt{\min(m,n)}},$$

The height and width of the input infrared image block are denoted as m and n respectively. The penalty parameter β in the augmented Lagrangian framework was chosen to be 0.025, and the multiplier μ was empirically set to 1.5.

For the proposed algorithm solving the FTVN model, we define the tolerance error as:

$$\mathrm{tol} = \frac{\left\|I - T^k - B^k - N^k\right\|_F}{\|I\|_F}, \tag{32}$$

where k indicates the current iteration index. Convergence is declared when $\mathrm{tol} < 10^{-5}$ or when the maximum number of iterations, $\mathrm{MaxIter} = 1000$, is reached.

Figs. 2-7 present detection outcomes across different methodologies alongside corresponding TDID plots, providing visual assessment of target enhancement and noise suppression capabilities. Ground-truth targets are circled in green for enhanced visibility. Analysis of Fig. 3 (Dataset 2) reveals minimal background clutter with two high-brightness targets. All approaches demonstrate effective detection, achieving perfect (100%) detection probabilities. For complex backgrounds (Figs. 4-7), most techniques successfully identify targets but exhibit substantial residual clutter in target images, leading to elevated false alarm rates. Notably, classical methods (IPIAPG, PSTNN, TV-PCP) produce widely distributed noise artifacts. While the proposed method shows minor residuals in datasets 4-6 (Figs. 5-7), its background residuals are significantly attenuated compared to alternatives, demonstrating distinct performance advantages in clutter suppression.

To quantitatively assess performance, we report the average SNRG and BSF for the proposed and comparative methods, summarized in Tables IV and V. In datasets 1, 4, and 5, the proposed method demonstrates excellent performance, achieving the highest average SNRG and BSF values, significantly outperforming other comparison methods. Specifically, in these datasets, the proposed method shows remarkable advantages in background noise suppression and target detection enhancement. In datasets 2 and 6, although the proposed method did not achieve the highest values, it still demonstrated strong competitiveness, ranking second in both cases. This indicates that the proposed method has strong robustness and versatility when processing datasets of varying complexity and can effectively improve detection performance across different scenarios.

D. Sensitivity Analysis

The proposed model incorporates several critical parameters requiring sensitivity analysis to optimize detection performance. Among these, the regularization parameters λ_1 and λ_2 exert significant influence on the optimization behavior governed by Eq.(11).

During the experiments, all other parameters are held constant, and only λ_1 and λ_2 are varied to evaluate their influence on performance. We let λ_1 take values in $\{0.01,$

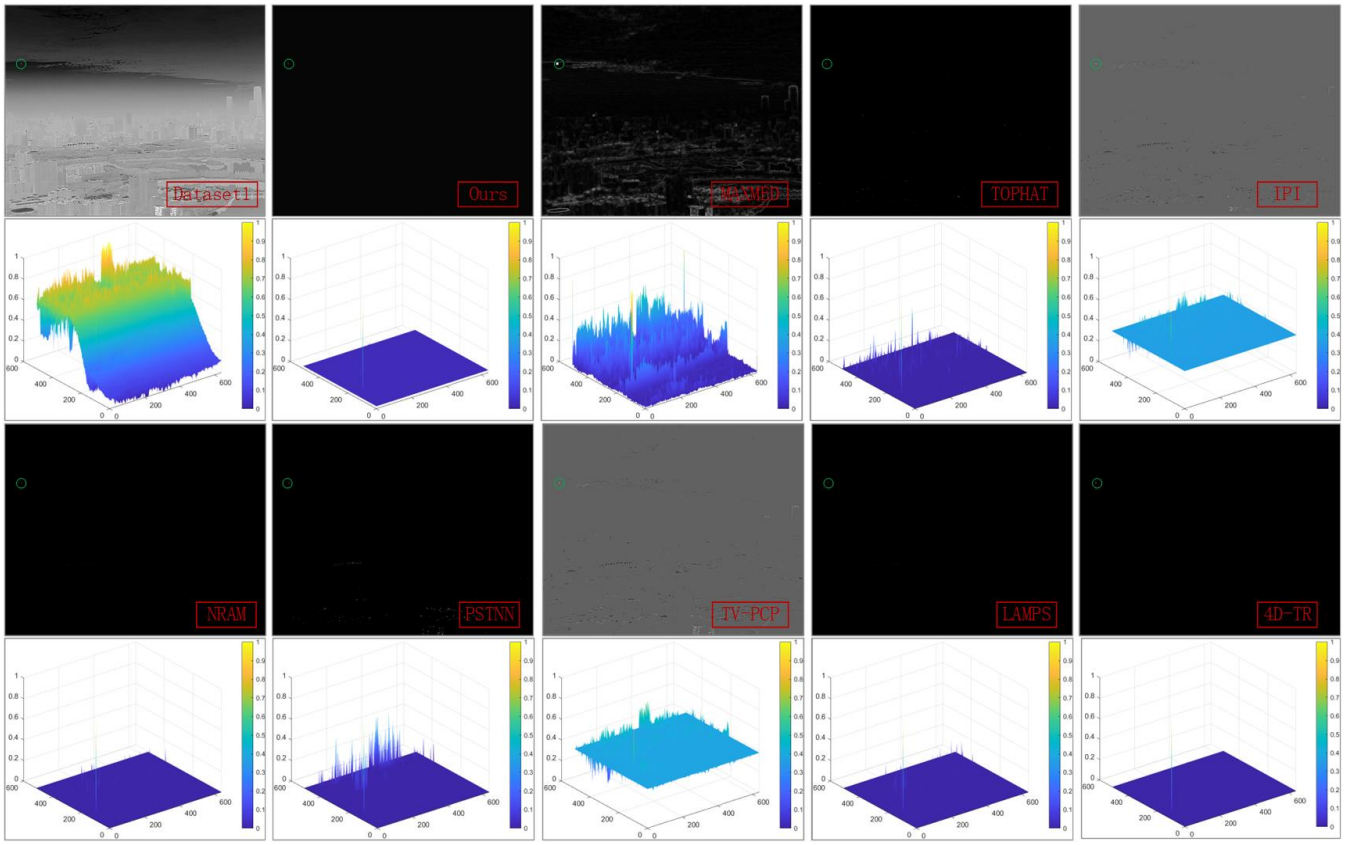

Fig. 2: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 1. Ground-truth targets are annotated with green circles.

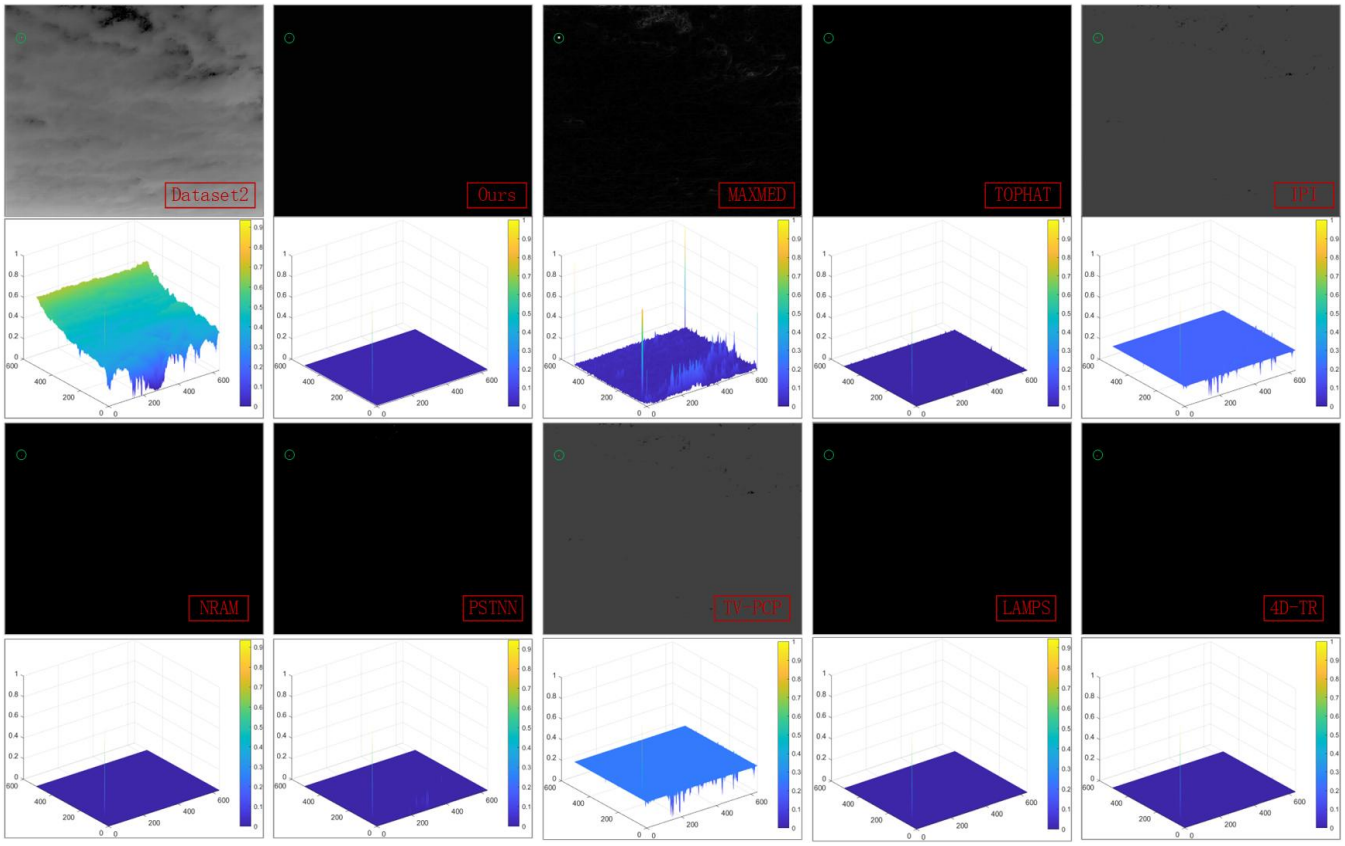

Fig. 3: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 2. Ground-truth targets are annotated with green circles.

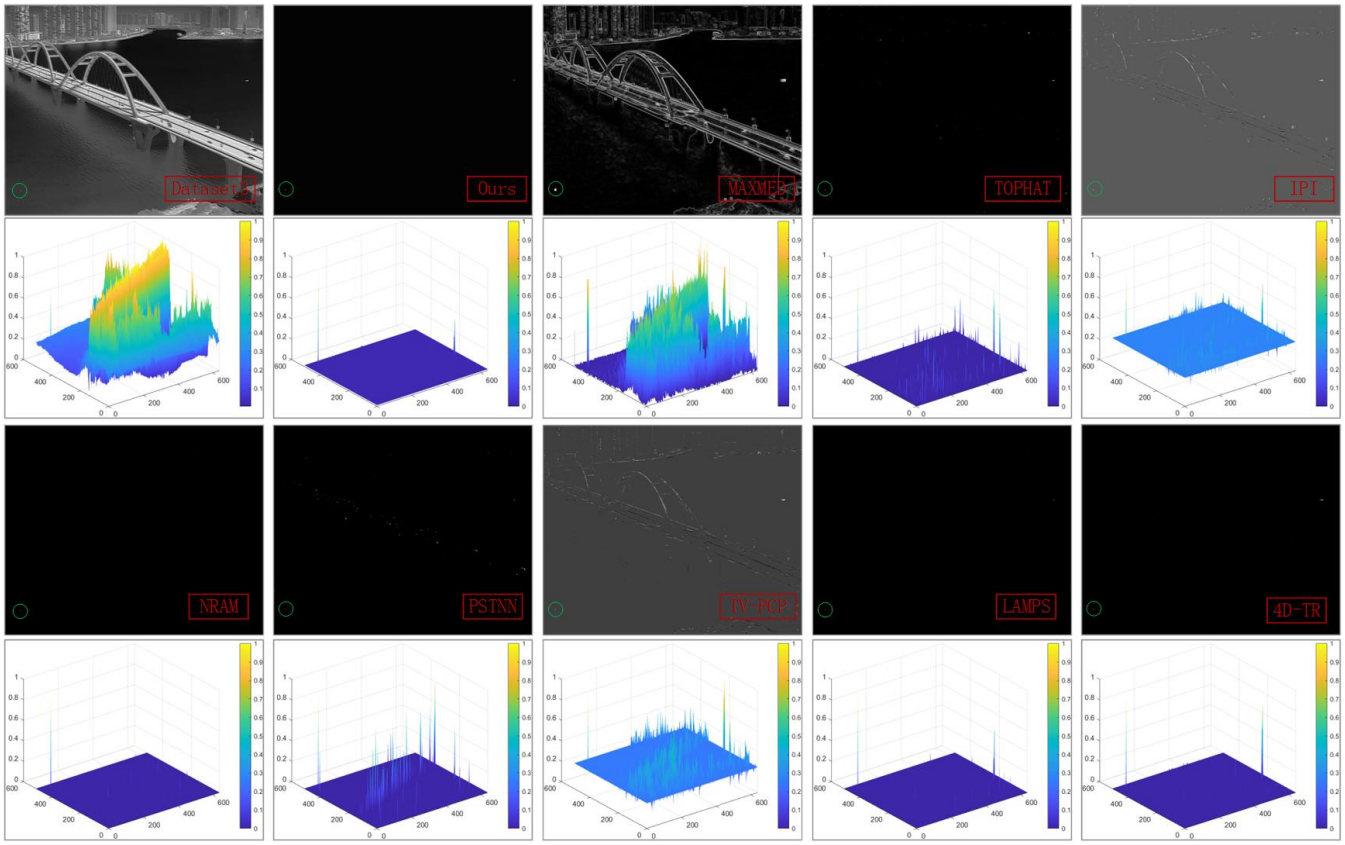

Fig. 4: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 3. Ground-truth targets are annotated with green circles.

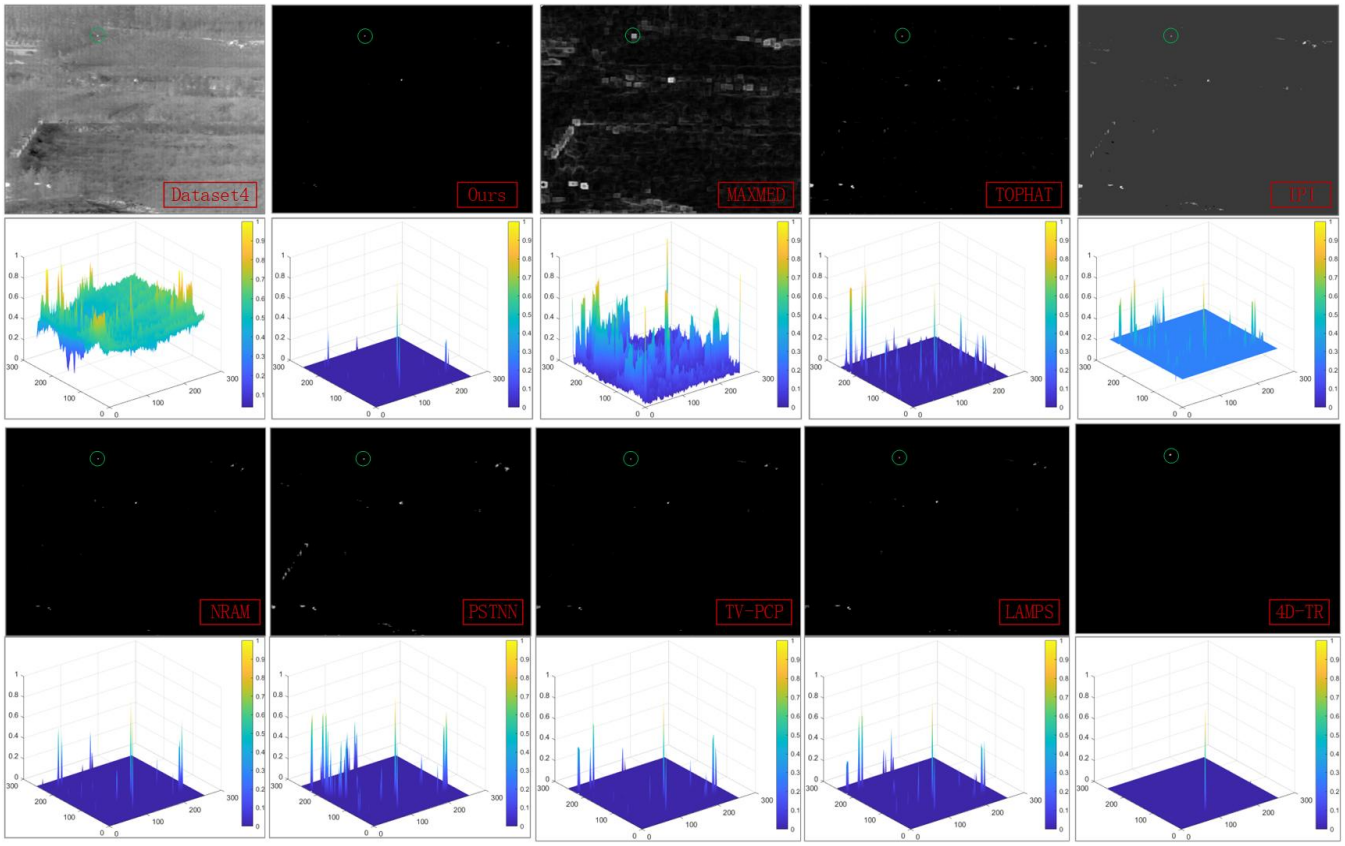

Fig. 5: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 4. Ground-truth targets are annotated with green circles.

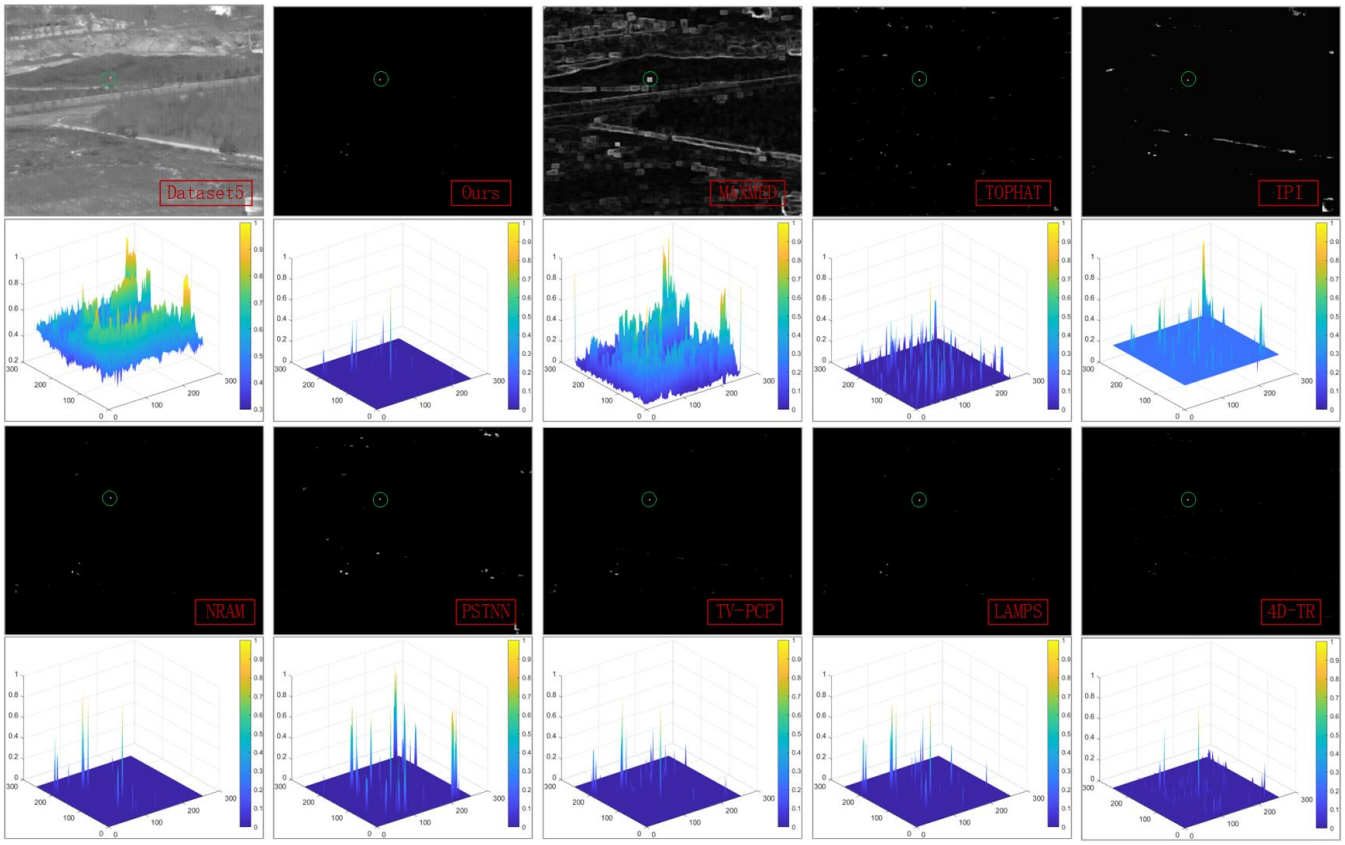

Fig. 6: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 5. Ground-truth targets are annotated with green circles.

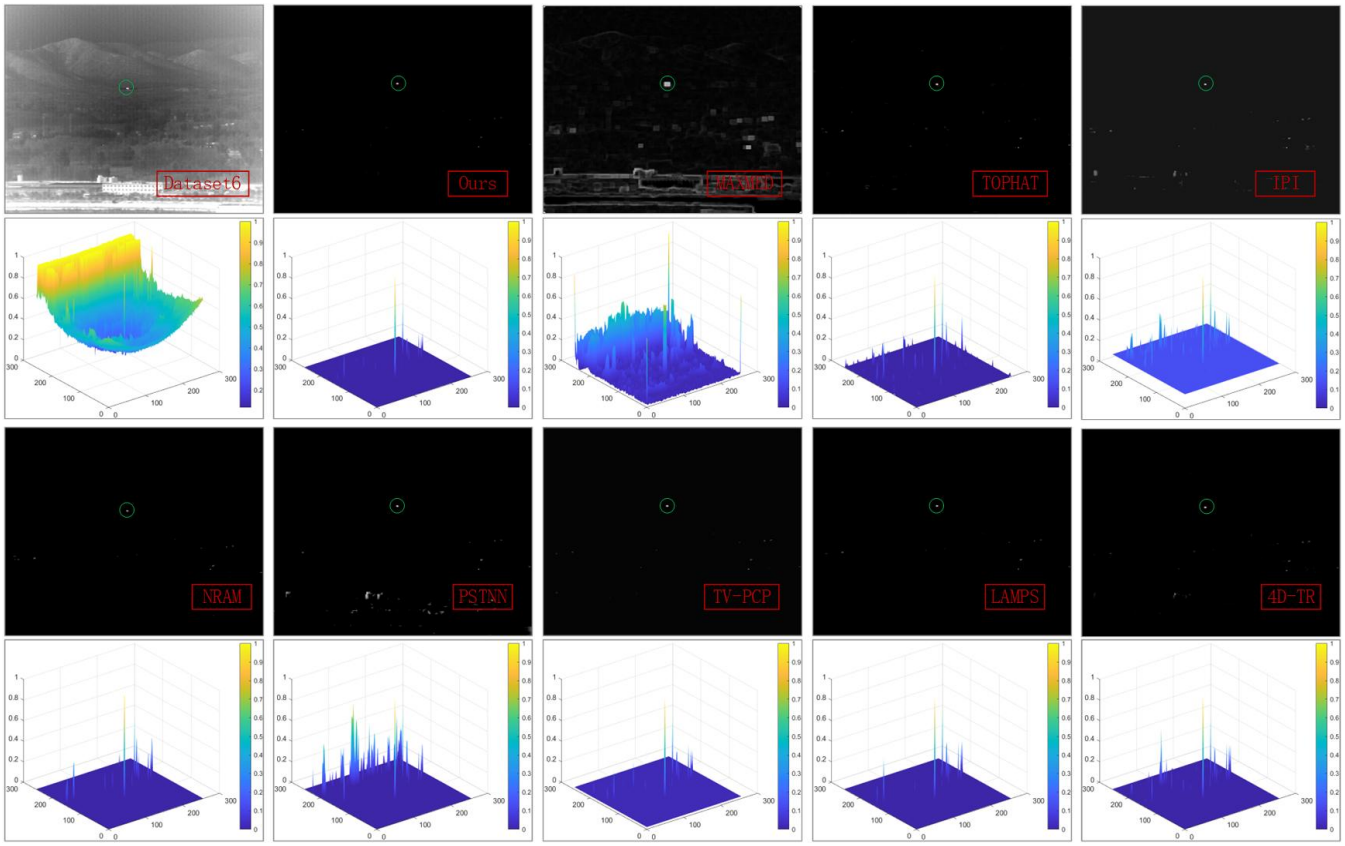

Fig. 7: Detection outcomes and TDID metrics across comparative methodologies for the initial frame of Dataset 6. Ground-truth targets are annotated with green circles.

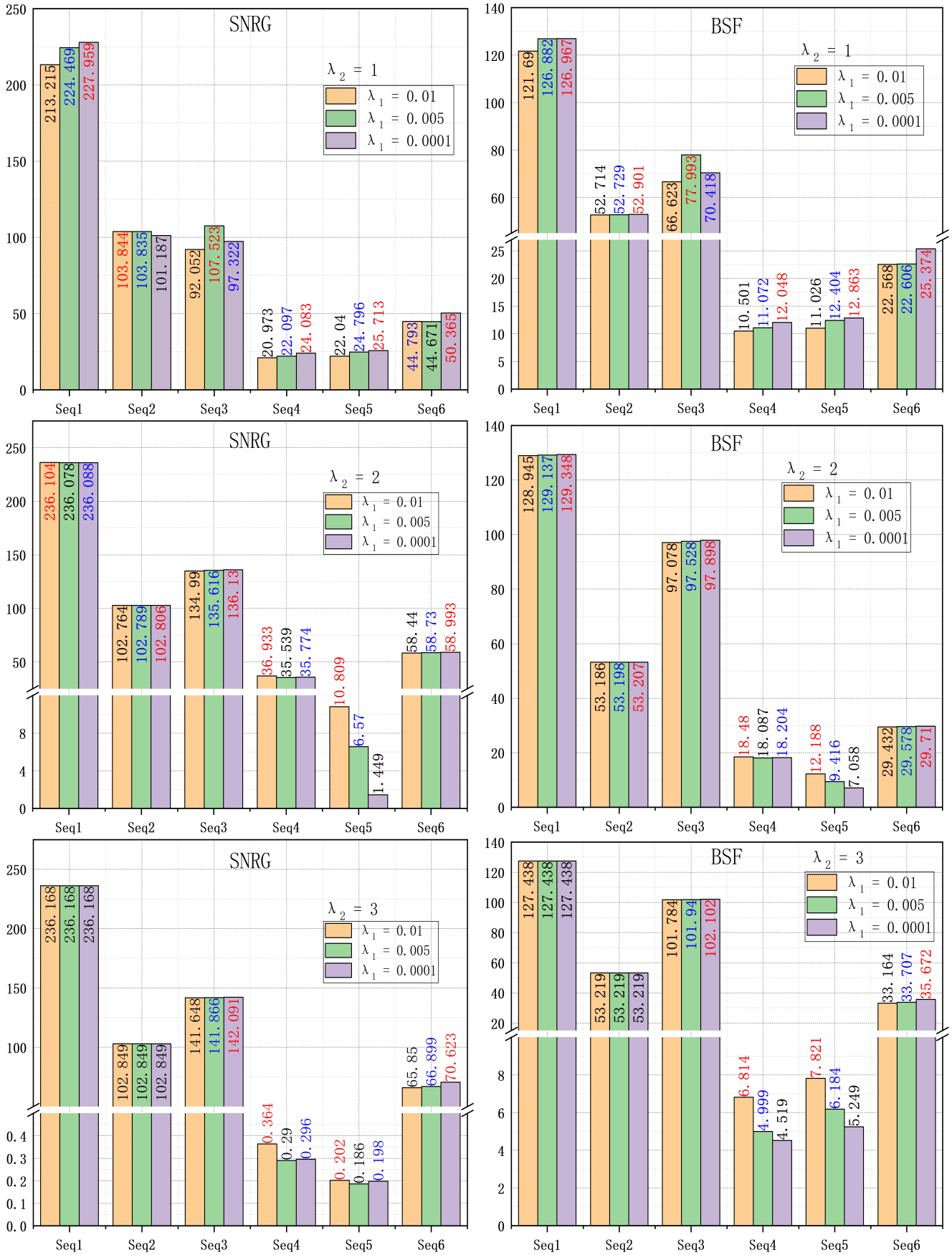

Fig. 8: The first column shows the impact of different combinations of values for λ_1 and λ_2 on SNRG, while the second column shows their impact on BSF. The values in the figure are all average values.

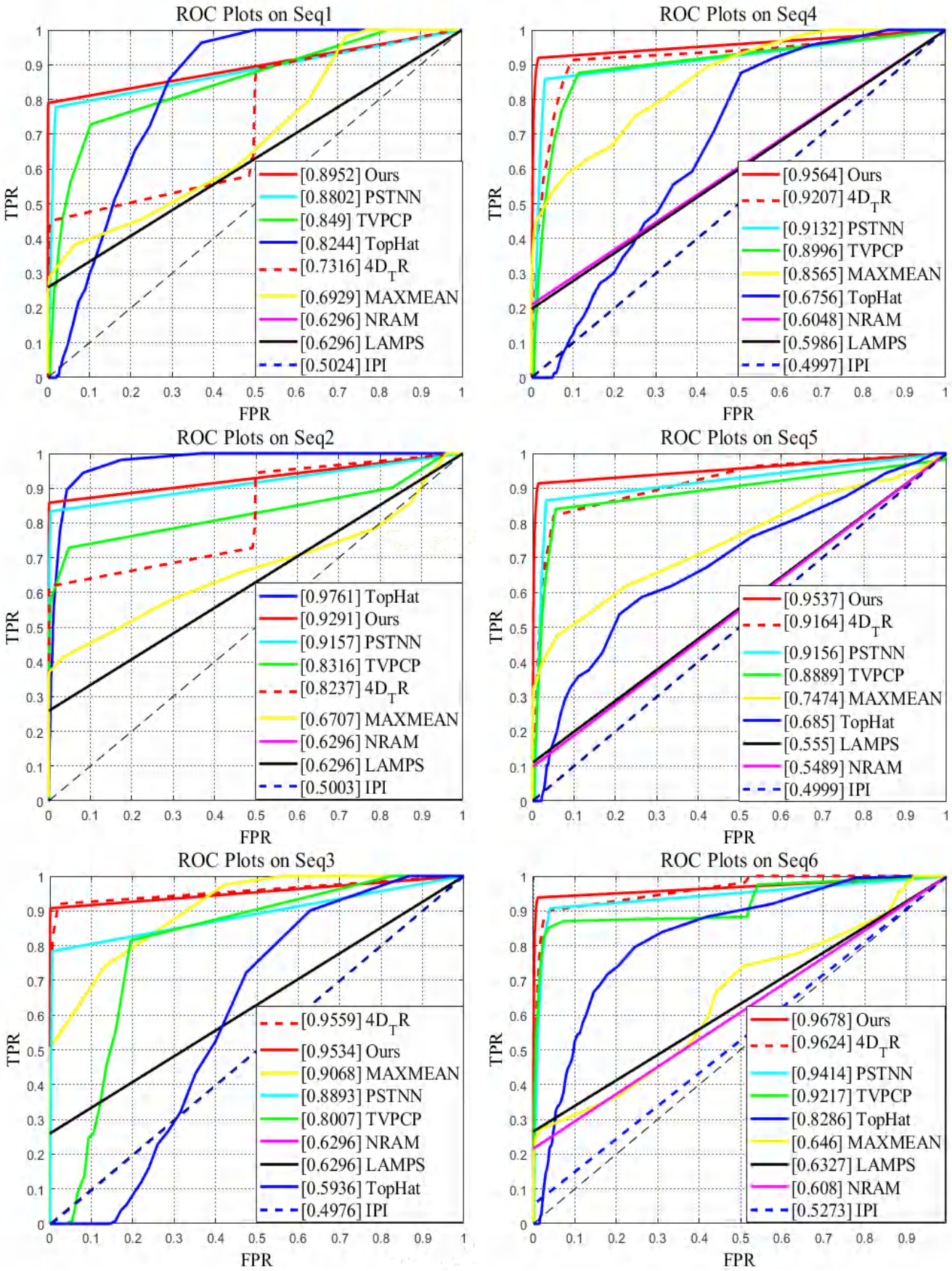

Fig. 9: The ROC curves presented in the figure provide a performance comparison of different approaches with respect to detection capability.

TABLE I: Overview of Six Representative Infrared Sequences.

Datasets	Background overview	Target overview
Dataset 1	Complex urban background. The background brightness varies greatly.	Small round target exhibiting low contrast relative to the nearby area.
Dataset 2	Bright cloud-covered background containing rich texture in specific sections.	A small round target, moving through the clouds.
Dataset 3	Ocean background. Complex artificial buildings near the coast.	A circular target with a bright dot. Moves slowly in sequence.
Dataset 4	Ground environment featuring significant clutter and bright background interference.	The target is an airplane with a random motion trajectory.
Dataset 5	Ground scene with heavy clutter.	An airplane observed at a long imaging distance, exhibiting significant size variation.
Dataset 6	Ground scene with heavy clutter. Complex artificial buildings near the mountains.	A single target on the move.

TABLE II: Key parameter configurations for the comparison approaches employed during experimentation.

Comparison Method	Acronyms	Key Configuration Parameters
Max-median Filter	MAXMED	Structural dimensions: 5×5
New-Top-hat Filter	TOPHAT	Structural Shape: disk; Size: 5×5
Partial Sum of Tensor Nuclear Norm	PSTNN	Patch size: 40×40; Sliding step: 40; $\lambda = 0.7/\sqrt{\max(n_1, n_2)} \times n_3$
Infrared Patch-Image with Accelerated Proximal Gradient Solver	IPIAPG	Patch size: 50×50; Sliding step: 20 $\lambda = 1/\sqrt{\max(m, n)}$
Total variation regularization and principal component pursuit	TV-PCP	Patch size: 50×50; Sliding step: 20 $\lambda_1 = 0.005$; $\lambda_2 = \frac{1}{\sqrt{\min(m,n)}}$; $\beta = 0.025$; $\gamma = 1.5$
Non-convex Rank Approximation Minimization	NRAM	Patch size: 50×50; Sliding step: 10; $\lambda = 1/\sqrt{\min(m, n)}$
Low-rank approximation and multiple sparse	LAMPS	Patch size: 50×50; Sliding step: 20 $\lambda = 0.4$; $\alpha = 0.01$; $\beta = 0.01$;
4-D Tensor Ring model	4D-TR	Patch size: 70×70; Tenmporal size: $N_3 = 15$; $L = 2$; $\lambda_1 = \sum_{i=1}^{l} \frac{L}{\sqrt{\max(\prod_{i=n}^{n+l-1} N_i, \prod_{i=n+l}^{n-1} N_i)}}$

0.005, 0.0001}, and λ_2 take values in {1, 2, 3} (here the value of λ_2 refers to the numerator in the equation $\lambda_2 = \frac{1}{\sqrt{\min(m,n)}}$, while the denominator remains unchanged) The average SNRG and BSF metrics under different settings of λ_1 and λ_2 are reported in Fig. 8. The results suggest several noteworthy observations. When the value of λ_1 is fixed, the average SNRG and BSF of most datasets show a slight increase as the value of λ_2 increases. However, when the value of λ_2 reaches 3, the average SNRG and BSF for datasets 4 and 5 drop sharply. When λ_2 is fixed, the average SNRG and BSF of most datasets show a slight increase as λ_1 decreases. In general, when λ_2 is set to 3, datasets 4 and 5 exhibit the worst detection performance, with a significant number of false alarms.

In Eq. (5), $D_x^\alpha x_{i,j}$ and $D_y^\alpha x_{i,j}$ denote order-α derivatives taken in the horizontal and vertical directions, respectively. where α represents the fractional order, typically within the range $(0, 1)$. The choice of α significantly affects the smoothness of gradient computation and denoising performance. A lower value α generally produces smoother gradients and reduces noise but may weaken the denoising effect. In contrast, a larger α value improves edge preservation but could not completely eliminate noise. We investigated the impact of different values α on the SNRG and BSF metrics of the proposed algorithm. As shown in Tab III, the key findings can be summarized as follows: for the SNRG metric, $\alpha = 0.4$ yields the best performance across all six datasets, with $\alpha = 0.2$ as the second-best choice (except for Dataset 5). For the BSF metric, $\alpha = 0.4$ also performs best (except for Dataset 2).

Next, the ROC curves for the nine methods evaluated on six image sequences are presented. These curves illustrate how the detection probability (Pd) varies with different false alarm rates (Fa), reflecting the trade-off between sensitivity and specificity. To further quantify performance, the Area Under the Curve (AUC) for each method is

TABLE III: BSF and SNRG indicators of each sequence under different α values

Datasets	$\alpha = 0.2$		$\alpha = 0.3$		$\alpha = 0.4$		$\alpha = 0.5$		$\alpha = 0.6$	
	BSF	SNRG	BSF	SNRG	BSF	SNRG	BSF	SNRG	BSF	SNRG
seq1	125.891	226.302	125.884	226.289	125.925	226.362	125.889	226.296	125.849	226.225
seq2	52.803	101.08	52.802	101.078	52.806	101.086	52.802	101.078	52.807	101.069
seq3	66.379	91.745	66.304	91.640	66.454	91.848	66.257	91.575	65.879	91.053
seq4	11.309	22.604	11.301	22.588	11.311	22.609	11.288	22.563	11.247	22.481
seq5	12.186	24.359	12.179	24.344	12.22	24.427	12.203	24.392	12.176	24.338
seq6	24.870	49.37 1	24.861	49.353	24.872	49.375	24.803	49.328	24.822	49.264

TABLE IV: Comparison of $\overline{SNRG}$ Across Methods for Datasets 1–6

Methods	$\overline{SNRG}$					
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
MAXMED	6.054	9.384	2.059	2.293	1.718	5.146
TOPHAT	59.607	97.101	40.837	9.4426	10.098	35.586
IPIAPG	22.803	40.931	29.977	8.299	12.118	24.794
TV-PCP	12.013	26.931	11.431	14.396	19.049	39.602
NRAM	181.234	104.063	130.491	14.3914	15.634	22.545
PSTNN	37.001	76.923	40.096	6.775	10.097	35.586
LAMPS	197.972	100.912	136.029	12.711	18.679	47.489
4D-TR	220.986	99.188	73.947	21.804	22.661	41.565
OURS	224.469	103.835	107.523	22.097	24.796	44.671

TABLE V: Comparison of $\overline{BSF}$ Across Methods for Datasets 1–6

Methods	$\overline{BSF}$					
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
MAXMED	3.453	4.779	1.591	1.239	0.974	2.714
TOPHAT	32.163	48.697	29.471	4.731	5.056	17.903
IPIAPG	25.002	25.968	31.421	5.309	7.564	15.287
TV-PCP	14.175	18.263	11.086	7.366	9.641	22.112
NRAM	97.761	56.886	93.817	7.198	7.821	22.437
PSTNN	15.331	41.609	28.936	3.396	5.056	17.903
LAMPS	106.792	52.488	97.704	6.991	9.347	24.302
4D-TR	124.578	51.041	53.386	10.919	11.337	21.927
OURS	126.882	52.729	77.993	11.072	12.404	22.606

calculated and annotated in the figure. Generally, a higher AUC indicates better detection capability. The ROC curves and corresponding AUC values of all methods on six sequences are shown in Figure 9. In most cases, the proposed method achieves the highest AUC, demonstrating superior overall detection by balancing Pd and Fa more effectively. The six sequences feature complex, non-uniform, and unsmooth backgrounds. The ROC curves confirm

that the proposed model maintains robust and effective performance in complex environments, highlighting its practical utility. Additionally, While ROC curves depict the PD–FA relationship, a thorough evaluation should also include precision, recall, and the F1 score. Across these measures, the proposed method consistently outperforms baseline approaches in detection accuracy.

V. Conclusions

To improve IPI-based detection performance in complex backgrounds, we propose a new infrared dim small target detection framework integrating fractional-order total variation regularization with a nonnegative L_1-norm constraint. This model effectively preserves the overall background pixel distribution while retaining detailed edges and corners. On six benchmark datasets, experimental evaluations indicate that the proposed method effectively reduces background clutter while enhancing target features.

References

[1] X. Mao and W.-h. Diao, "Criterion to evaluate the quality of infrared small target images," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 30, pp. 56–64, 2009.

[2] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," in *Processing of SPIE 1954*, 1993.

[3] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," in *Processing of SPIE 3809*, 1999.

[4] X. Shao, H. Fan, G. Lu, and J. Xu, "An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system," *Infrared Physics & Technology*, vol. 55, no. 5, pp. 403–408, 2012.

[5] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 574–581, 2013.

[6] X. Cao, C. Rong, and X. Bai, "Infrared small target detection based on derivative dissimilarity measure," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 3101–3116, 2019.

[7] S. Pan, S. Zhang, M. Zhao, and B. An, "Infrared small target detection based on double-layer local contrast measure," *Acta Photonica Sinica*, vol. 49, p. 0110003, 2020.

[8] P. Du and A. Hamdulla, "Infrared moving small-target detection using spatial–temporal local difference measure," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 10, pp. 1817–1821, 2019.

[9] S. Xu, S. Zheng, W. Xu, R. Xu, C. Wang, J. Zhang, X. Teng, A. Li, and L. Guo, "Hcf-net: Hierarchical context fusion network for infrared small object detection," in *IEEE International Conference on Multimedia and Expo*, 2024.

[10] H. Yang, T. Mu, Z. Dong, Z. Zhang, B. Wang, W. Ke, Q. Yang, and Z. He, "Pbt: Progressive background-aware transformer for infrared small target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.

[11] C. Bao, J. Cao, Y. Ning, T. Zhao, Z. Li, Z. Wang, L. Zhang, and Q. Hao, "Improved dense nested attention network based on transformer for infrared small target detection," *arXiv preprint arXiv:2311.08747*, 2023.

[12] S. Li and W. Liu, "Small target detection model in aerial images based on yolov7x+," *Engineering Letters*, vol. 32, pp. 436–443, 2024.

[13] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, pp. 1–37, 2011.

[14] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," *Advances in Neural Information Processing Systems*, vol. 22, 2009.

[15] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4996–5009, 2013.

[16] Y. Dai, Y. Wu, and Y. Song, "Infrared small target and background separation via column-wise weighted robust principal component analysis," *Infrared Physics & Technology*, vol. 77, pp. 421–430, 2016.

[17] J. Guo, Y. Wu, and Y. Dai, "Small target detection based on reweighted infrared patch-image model," *IET Image Processing*, vol. 12, no. 1, pp. 70–79, 2018.

[18] Z. Kang, C. Peng, and Q. Cheng, "Robust pca via nonconvex rank approximation," in *IEEE International Conference on Data Mining*, 2015, pp. 211–220.

[19] F. Zhou, Y. Wu, Y. Dai, and P. Wang, "Detection of small target using schatten 1/2 quasi-norm regularization with reweighted sparse enhancement in complex infrared scenes," *Remote Sensing*, vol. 11, no. 17, p. 2058, 2019.

[20] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint l 2, 1 norm," *Remote Sensing*, vol. 10, no. 11, p. 1821, 2018.

[21] X.-R. Feng, H.-C. Li, J. Li, Q. Du, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using sparsity-constrained deep nonnegative matrix factorization with total variation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 6245–6257, 2018.

[22] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[23] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image and Vision Computing*, vol. 63, pp. 1–9, 2017.

[24] R. Wang, H.-C. Li, A. Pizurica, J. Li, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using double reweighted sparse regression and total variation," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1146–1150, 2017.

[25] S. S. Rawat, S. K. Verma, and Y. Kumar, "Reweighted infrared patch image model for small target detection based on non-convex p-norm minimisation and tv regularisation," *IET image processing*, vol. 14, no. 9, pp. 1937–1947, 2020.

[26] Y. Sun, J. Yang, and W. An, "Infrared small target detection based on reweighted infrared patch-image model and total variation regularization," in *Processing of SPIE 11155*, 2019.

[27] W. Li, M. Zhao, X. Deng, L. Li, L. Li, and W. Zhang, "Infrared small target detection using local and nonlocal spatial information," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3677–3689, 2019.

[28] J. Gao, Y. Guo, Z. Lin, and W. An, "Infrared small target detection using multiscale gray and variance difference," in *Chinese Conference on Pattern Recognition and Computer Vision*, 2018, pp. 53–64.

[29] J. Han, C. Liu, Y. Liu, Z. Luo, X. Zhang, and Q. Niu, "Infrared small target detection utilizing the enhanced closest-mean background estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 645–662, 2020.

[30] D. Liu, L. Cao, Z. Li, T. Liu, and P. Che, "Infrared small target detection based on flux density and direction diversity in gradient vector field," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 7, pp. 2528–2554, 2018.

[31] P. Yang, L. Dong, and W. Xu, "Infrared small maritime target detection based on integrated target saliency measure," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2369–2386, 2021.

[32] D. Chen, Y. Chen, and D. Xue, "Fractional-order total variation image restoration based on primal-dual algorithm," *Abstract and Applied Analysis*, Article ID 585310, 2013.

[33] Y.-F. Pu, J.-L. Zhou, P. Siarry, N. Zhang, and Y.-G. Liu, "Fractional partial differential equation: Fractional total variation and fractional steepest descent approach-based multiscale denoising model for texture image," *Abstract and Applied Analysis*, Article ID 483791, 2013.

[34] Z. Wang, H. Li, Q. Ling, and W. Li, "Robust temporal-spatial decomposition and its applications in video processing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 3, pp. 387–400, 2012.

[35] G. Chen, J. Zhang, D. Li, and H. Chen, "Robust kronecker product video denoising based on fractional-order total variation model," *Signal Processing*, vol. 119, pp. 1–20, 2016.

[36] P. D. Tao and L. H. An, "Convex analysis approach to dc programming: theory, algorithms and applications," *Acta Mathematica Vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.

[37] C. Li, *An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing*. Rice University, 2010.