



OPEN ACCESS

EDITED BY

Yimian Dai,
Nanjing University of Science and
Technology, China

REVIEWED BY

Fei Zhou,
Henan University of Technology, China
Renke Kou,
Air Force Engineering University, China

*CORRESPONDENCE

Jinhui Han

✉ hanjinhui@zkn.edu.cn

RECEIVED 27 February 2025

ACCEPTED 09 April 2025

PUBLISHED 08 May 2025

CITATION

Han J, Moradi S, Wang W, Li N, Zhao Q and
Luo Z (2025) A hybrid contrast method for
infrared small dim target detection.
Front. Mar. Sci. 12:1584345.
doi: 10.3389/fmars.2025.1584345

COPYRIGHT

© 2025 Han, Moradi, Wang, Li, Zhao and Luo.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

A hybrid contrast method for infrared small dim target detection

Jinhui Han^{1*}, Saed Moradi², Wei Wang³, Nan Li¹, Qian Zhao¹
and Zhen Luo³

¹College of Physics and Telecommunication Engineering, Zhoukou Normal University, Zhoukou, China, ²Department of Electrical Engineering, University of Isfahan, Isfahan, Iran, ³School of Artificial Intelligence, Zhoukou Normal University, Zhoukou, China

Infrared (IR) small dim target detection under complex background is crucial in many fields, such as maritime search and rescue. However, due to the interference of high brightness background, complex edges/corners and random noises, it is always a difficult task. Especially, when a target approaches a high brightness background area, the target will be easily submerged. In this paper, a new contrast method framework named hybrid contrast measure (HCM) is proposed, it consists of two main modules: the relative global contrast measure (RGCM) calculation, and the small patch local contrast weighting function. In the first module, instead of using some neighboring pixels as benchmark directly during contrast calculation, the sparse and low rank decomposition method is adopted to get the global background of a raw image as benchmark, and a local max dilation (LMD) operation is applied on the global background to recover edge/corner information. A Gaussian matched filtering operation is applied on the raw image to suppress noises, and the RGCM will be calculated between the filtered image and the benchmark to enhance true small dim target and eliminate flat background area simultaneously. In the second module, the Difference of Gaussians (DoG) filtering is adopted and improved as the weighting function. Since the benchmark in the first module is obtained globally rather than locally, and the patch size in the second module is very small, the proposed algorithm can avoid the problem of the targets approaching high brightness backgrounds and being submerged by them. Experiments on 14 real IR sequences and one single frame dataset show the effectiveness of the proposed algorithm, it can usually achieve better detection performance compared to the baseline algorithms from both target enhancement and background suppression point of views.

KEYWORDS

remote sensing, infrared small dim target detection, relative global contrast measure, local contrast measure, hybrid contrast measure, sparse and low rank decomposition

1 Introduction

In the field of guidance, early warning, airborne or spaceborne monitoring/surveillance, maritime search and rescue, etc., infrared detection system has become an effective supplement or alternative for traditional visible light and radar detection systems (Kou et al., 2023). However, in some practical applications, targets are far from the detector and occupy only a few pixels (usually less than 9×9) in the output image with low intensity value (Pang et al., 2022a), which are usually called IR small targets or IR small dim targets. The detection of IR small dim targets is highly challenging (Luo et al., 2024; Ma et al., 2022), as shown in Figure 1: Firstly, the small size of the target results in a lack of significant shape or texture information; Secondly, the low intensity value makes it difficult to obtain the target directly in the raw image; Thirdly, there are usually various complex backgrounds such as buildings, clouds and sea waves in real applications, they may have high brightness and complex edges, resulting many false alarms; Finally, some bad detector pixels and random electrical noise may cause some pixel-size noise with high brightness (PNHB), they may also bring false alarms.

So far, a lot of methods have been proposed to address the problem of IR small dim target detection under complex backgrounds, which can be mainly divided into two categories: data-driven methods and model-driven methods (Kou et al., 2023). The data-driven methods typically use prepared image data to train a deep convolution network, enabling the network to classify the input data as targets or backgrounds, they can be further divided into the two-stage methods such as RCNN (Girshick et al., 2014; Ren et al., 2017), the single stage methods such as YOLO (Redmon et al., 2016; Liu et al., 2024; Xu et al., 2024), and the deep unfolding methods such as RPCANet (Wu et al., 2024). However, data-driven methods only work well when the input data and the training data have the same distribution function, but it may not be always true in practical applications. In addition, a deep convolution network usually contains a large number of parameters, which is difficult to train and deploy on a single-chip system.

The model-driven methods typically design a handy-crafted feature model according to the feature difference between the true small dim target and the complex background firstly, then search for targets in the feature saliency map. Compared to the data-driven methods, the model-driven methods are easier to understand and implement for real applications. Therefore, the model-driven methods still attract a lot of attention nowadays.

The feature model is a key module in model-driven methods, which can be designed between consecutive frames or within a single frame. These methods are usually called as sequence-based methods and frame-based methods, respectively (Luo et al., 2022). The sequence-based methods (Zhang et al., 2021; Du and Hamdulla, 2020a; Dang et al., 2023; Pang et al., 2022b) utilize information from multiple frames simultaneously for target detection, therefore they usually have better performance. However, such algorithms typically have higher computational costs and require more processing resources as they need to consider a large amount of information. On the contrary, the frame-based algorithms perform target detection within a single frame, so they usually require less computation and storage space, making them easier to implement. Moreover, a frame-based algorithm is usually adopted as a basic module in some sequence-based algorithms. Therefore, in this paper we focus on the frame-based IR small target detection.

According to the different information used during feature extraction, existing frame-based model-driven algorithms can be further divided into background estimation methods, morphological methods, directional derivative/gradient methods, local contrast methods, frequency filter methods, sparse representation methods, and sparse and low rank decomposition methods.

1.1 Background estimation methods

Background estimation methods estimate the background value of each pixel using its neighboring benchmark, then subtract the background image from the raw image to get the foreground image.

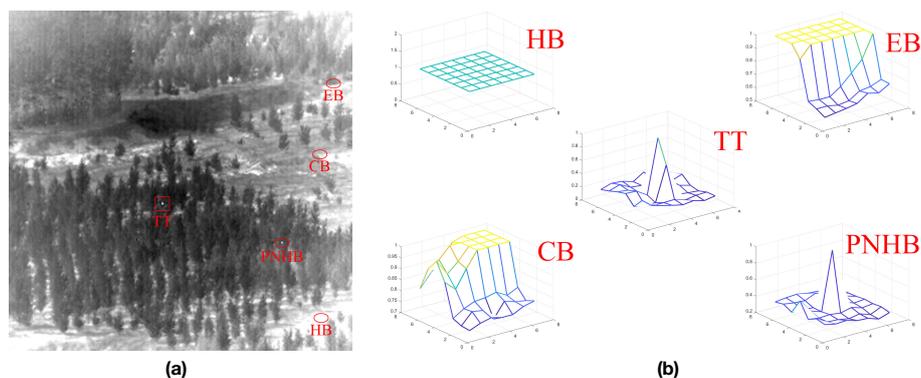


FIGURE 1

Distributions of real IR small target (labeled with rectangle) and common interference (labeled with ellipses). (a) A raw IR image; (b) Distributions of different components, TT represents true small target, HB represents high-brightness background, EB represents background edge, CB represents background corner, and PNHB represents pixel-size noise with high brightness.

Since the gray value of each pixel in an IR image consists of the background value, the target value and the random noise, we will be able to obtain the target easily in the foreground. Current background estimation methods include median filtering (Yang and Shen, 1998), max mean/max medium filtering (Deshpande et al., 1999), and some adaptive background estimation methods such as Two Dimensional Least Mean Square (TDLMS) (Ding and Zhao, 2015), etc.

1.2 Morphological methods

Morphological methods firstly design a specific shaped structural window according to the characteristics of the IR small dim targets, then use this window to do some morphological operations such as erosion and dilation at each position of the raw image to highlight the target and suppress background and noise. The ring-like new Top Hat structure window (Bai and Zhou, 2010) is an excellent representative of morphological methods for IR small dim target detection, which was further developed by Deng et al. (2021) and Zhu et al. (2020a); Zhang et al. (2024) designed two dilate structures to enhance true target and suppress clutters, respectively; Peng et al. (2024) designed a dual structure template with eight directions to distinguish between background edges and real targets; etc.

1.3 Directional derivative/gradient methods

Directional derivative/gradient methods take the gray value of a pixel as a scalar. If we calculate the directional derivative/gradient of a pixel, we can get directional information. Therefore, the number of features used to distinguish between the target and background will increase during this process, and the detection performance will be better. For example, Lu et al. (2022a) used a small kernel model to calculate directional derivatives; Bi et al. (2020) proposed the high-order directional derivatives; Yang et al. (2022) adopted the multidirectional difference measure as a weighting function of the multidirectional gradient; Hao et al. (2024) designed a gradient method that can adaptively handle multiscale targets; etc.

1.4 Local contrast methods

Local contrast methods are bionic methods inspired by the contrast mechanism of the human visual system. A human can quickly and accurately capture a target which is locally salient in the image, while large-area high brightness backgrounds will be ignored, because the human eyes are more sensitive to local contrast information rather than the brightness information (Itti et al., 1998). Therefore, using local contrast instead of brightness as the basis for IR small dim target detection, better performance can be achieved.

The core of the local contrast is the dissimilarity between a current pixel and its neighboring benchmark. According to the size

of the benchmark, existing local contrast methods can be divided into large patch methods and small patch methods.

1.4.1 Large patch local contrast methods

Large patch methods first design a double-layer or tri-layer window in which the inner layer is used to capture target and the outer layer is used to capture neighboring benchmark, then slide the window on the whole image and calculate the local contrast information at each pixel. For example, Chen et al. (2014) proposed a Local Contrast Measure (LCM) algorithm based on a double-layer nested window, it takes the ratio of the central cell and the surrounding cell as local contrast; Han et al. (2014) proposed an improved LCM (ILCM) algorithm that introduces the average of the central cell to suppress noise; Qin and Li (2016) proposed a novel LCM (NLCM) algorithm that only averages some largest pixels of each cell to protect true targets; Han et al. (2018) proposed the relative LCM (RLCM) in which both ratio and difference operations are used to enhance target and suppress clutter simultaneously; Wei et al. (2016) proposed a multiscale patch-based contrast measure (MPCM) to deal with targets of unknown size; Han et al. (2020) and Han et al. (2021) designed two tri-layer windows to deal with targets of unknown size only using single scale calculation, which was further developed by Liu et al. (2023); Han et al. (2022) extended the gray contrast to feature contrast; etc. In these methods, a large contrast value can usually be extracted when the target is locally salient, because its benchmark is the surrounding background of the target. However, it also has disadvantages: if the background in the field of view is complex and the target is not locally salient, that is, when the target approaches a high brightness background, the high brightness background may be probably included in the patch window and selected as the benchmark, and the true target will be submerged.

1.4.2 Small patch local contrast methods

The reason for the drawback of the large patch methods is that the patch size in them is too large, for example, in ILCM, the patch size can reach up to 24×24; in NLCM, it reaches up to 30×30. A natural idea to address this drawback is to reduce the size of the patch window. Researchers find that a true IR small dim target usually attenuated uniformly in all directions, which means the gray value of its center is usually larger than its surroundings. Therefore, we can still obtain the local contrast information of a true target even if the patch size is smaller than the size of the target. For example, Shao et al. (2012) used the Laplacian of Gaussian (LoG) filter template with positive center coefficients and negative surrounding coefficients for contrast calculation; Wang et al. (2012) proposed a similar but simpler filter template named Difference of Gaussian (DoG) filter; Han et al. (2016) used elliptical Gabor functions instead of the circular Gaussian functions as kernel functions to distinguish true targets and complex background edges; Chen et al. (2023a) and Guan et al. (2020a) used LoG and Gaussian filter as preprocessing stage; etc. In these small patch algorithms, the template size is usually set to only 5×5, so that the risk of the target being submerged will be reduced as much as possible. However, if the target is large enough, the

extracted contrast value will be small, because its benchmark still belongs to target pixels.

Due to their advantages such as theoretical simplicity, ease of implementation, and ability to enhance true targets and suppress complex backgrounds simultaneously, local contrast methods still attract a lot of attention currently. However, the large patch methods cannot work well when the target approaches a high brightness background, and the small patch methods cannot work well when the target is relatively large. To improve the detection performance further, some researchers also combine the local contrast with other methods. For example, Cui et al. (2016) combined local contrast with the support vector machines; Deng et al. (2016); Deng et al. (2017) and Yao et al. (2023) introduced local entropy to weight local contrast; Du and Hamdulla (2020b) used local smoothness as weighting function; Xiong et al. (2021) and Zhou and Wang (2023) combined with local gradient; Kou et al. (2022) combined with density peak; Wang et al. (2024) used the variance as weighting function; Wei et al. (2022) used facet filtering as preprocessing stage; Han et al. (2019) used TDLMS to get the background benchmark; Tang et al. (2023) used a dilation operator with a ring structure to get background benchmark; Han et al. (2021) selected the surrounding gray value closest to the center in eight directions as the benchmark; etc. However, these methods can only alleviate rather than solve these problems.

1.5 Frequency filter methods

Frequency filter methods assume that different components in the raw image occupy different bands in frequency domain: the background occupies the low-frequency band since background is usually flat and continuously distributed in large area; the target occupies the high-frequency band since it has a significant discontinuity with its neighborhood; the noise occupies the highest frequency band since noise is usually random. If we first transform the original image into the frequency domain and use a frequency filter, the target will be detected. For example, Yang et al. (2004) used Butterworth high-pass filters to obtain true targets; Qi et al. (2014) proposed a detection method based on quaternion Fourier transform; Gregoris et al. (1994) used the wavelet transform to get the frequency information of the raw image; Kong et al. (2016) used Haar wavelet to detect the sea-sky line first, and then detected targets in the next steps; Chen et al. (2019) combined local contrast with some frequency domain algorithms; etc. However, if the backgrounds are very complex, some edges and corners also contain a lot of high-frequency information, making it difficult to distinguish them from true targets.

1.6 Sparse representation methods

The main idea of sparse representation is to use an overcomplete dictionary containing many atoms to linearly represent the original data. If some atoms have a high correlation with the original data, their corresponding coefficients will be

relatively large, so the coefficient vector will present with obvious sparsity. Based on this, Zhao et al. (2011) used many simulated IR small targets as atoms to construct an overcomplete dictionary; He et al. (2015) discussed the feature of background and found that the background is usually low rank, so they proposed a representation method with low rank constraints; Zhang et al. (2017) first used a particle filter to construct a saliency map, and then represented the saliency map instead of the raw image; Qin et al. (2016) and Liu et al. (2017) constructed two dictionaries, one is used to represent targets and the other is used to represent backgrounds; Chen et al. (2023b) utilized an adaptive group sparse representation method to denoise the raw IR image; etc. However, the shapes of the target and the background are ever changing between different practical applications, so it is extremely difficult to construct a dictionary that covers all situations.

1.7 Sparse and low rank decomposition methods

Sparse and low rank decomposition methods consider that a raw image consists of two parts: the background image, and the foreground image. The background in nature usually has a certain degree of non-local self-similarity, so the rank of the background image is relatively low; The foreground image mainly contains small targets and random noises, so it is usually very sparse. If we decompose the raw image into a matrix with low rank and a matrix with sparsity, we will be able to identify the targets easily in the sparse matrix. According to the type of the original data used for decomposition, the sparse and low rank decomposition methods can be further divided into two main categories: the image patch type, and the tensor type.

1.7.1 Image patch type sparse and low rank decomposition methods

Image patch type sparse and low rank decomposition methods first construct an image patch matrix, then decompose it by iteration. For example, Gao et al. (2013) firstly divide the original image into many image patches, then stretch them into column vectors and form all vectors as a new matrix, named infrared patch image (IPI), then used Robust Principal Component Analysis (RPCA) algorithm to decompose it; Dai et al. (2017) only focus on some small singular values during RPCA decomposition to maintain complex background edges; Wang et al. (2017); Fang et al. (2020) and Zhu et al. (2020b) embedded some regularization factors in the objective function of RPCA decomposition to constrain the results; Hao et al. (2023) proposed a novel continuation strategy based on the proximal gradient algorithm to suppress strong edges; etc.

1.7.2 Tensor type sparse and low rank decomposition methods

Tensor type sparse and low rank decomposition methods first stack image blocks into a three-dimensional tensor, then decompose it by iteration. For example, Dai and Wu (2017)

constructed the infrared patch tensor model, besides, some local priors in the image are used as weighting function, too; Zhang et al. (2019) simplified Dai's model and abandoned the weighting function; Fan et al. (2022) proposed a new anisotropic background feature as the weighting function for the infrared patch tensor model; Guan et al. (2020b) used global sparsity information as a weighting function for local contrast; Lu et al. (2022b) utilized some gradient information as weighting function during sparse and low rank decomposition; Zhang et al. (2023) utilized the local entropy as weighting function; etc.

Generally speaking, the sparse and low rank decomposition methods directly extract global target features from the whole IR image, so they can achieve good detection performance, especially when the target is not locally salient, for example, if a target approaches high brightness background. However, current algorithms only focus on the sparse foreground and attempt to directly find targets within it, but when the background in the field of view is very complex, a few background edges/corners and noise components also have sparse features and will be easily decomposed into the foreground, seriously interfering with the detection of real targets.

1.8 The main work of this paper

As mentioned earlier, when the target is locally salient, the large patch local contrast methods, such as LCM, ILCM, etc., can extract contrast information to the maximum extent, but they will fail when the target is not locally salient. This is because they directly utilize some neighboring pixels as the benchmark for a current pixel when calculating its contrast information. Therefore, we studied how to improve the benchmark selection principle to address this issue. We found that the sparse and low rank decomposition methods can decompose a raw image into a low rank background matrix and a sparse foreground matrix, and, since this decomposition is performed at the global level, the background at a true target's position will be more accurate and not affected by neighboring high brightness area. Thus, we propose that the decomposed low rank background matrix will be more suitable as the benchmark for calculating contrast information, and a new detection framework named hybrid contrast measure (HCM) for detecting IR small dim target in complex background is proposed in this paper.

The main work and contributions of this paper can be described as:

1. A new detection framework named HCM is proposed, it consists of two main modules: the relative global contrast calculation, and the small patch local contrast weighting function. Both global and local information are utilized in this new framework to achieve better detection performance even when the background is very complex and the target is not locally salient.
2. In the relative global contrast calculation module, the IPI model and the sparse and low rank decomposition are

adopted to get the global background as the benchmark for contrast calculation. However, due to the sparsity of edge and corner information, it is difficult to maintain them in a low rank global background. In this paper, we analyzed the essence of matrix low rank conversion and stated that some local operations can also be introduced on the global background to recover edge and corner information. Therefore, a simple but effective local max dilation (LMD) method is proposed and used on the estimated background image.

3. Inspired by our former work in the field of local contrast, we propose the relative global contrast measure (RGCM) between a raw image and its global benchmark to enhance true small dim target and eliminate flat background area simultaneously. Especially, before global contrast calculation, a Gaussian filter is applied on the raw image to suppress PNH better.
4. In the small patch local contrast weighting function module, the DoG filter is utilized to obtain local contrast information, which is then used as a weighting function after a simple non-negative operation to get the final saliency map. The advantage of choosing DoG is that its template is small enough, so we don't have to worry about the problem of the targets approaching high brightness backgrounds and being submerged by them.

2 Methodology

The framework of proposed HCM method is shown in Figure 2, there are two main modules in it: the RGCM calculation, and the small patch local contrast weighting function.

2.1 RGCM calculation

2.1.1 Global background separation

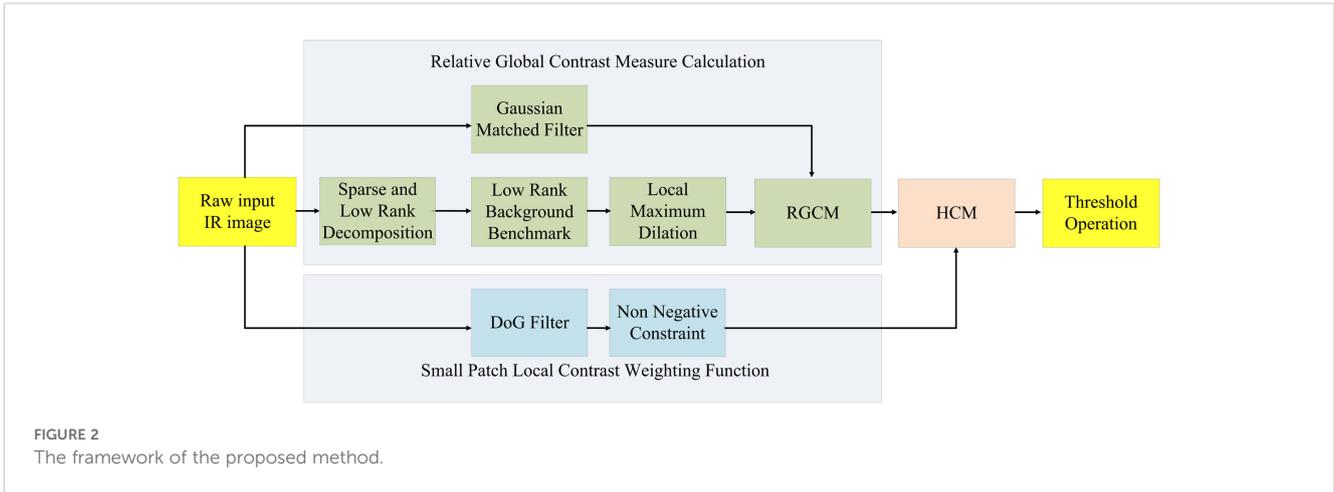
A raw IR image is usually modeled as the summation of three components, the background image, the target image, and the noise image, as shown in Equation 1:

$$I(x, y) = I_B(x, y) + I_T(x, y) + I_N(x, y) \quad (1)$$

where (x, y) is the coordinate of each pixel in the image, I is the raw image, I_B is the background image, I_T is the target image, and I_N is the noise image.

In practical applications, the background is usually self-similar, so I_B is usually a low rank image. The size of small target is usually very small, so I_T is usually a sparse image. If we use a sparse and low rank decomposition algorithm such as RPCA to divide a raw IR image into a sparse image and a low rank image, the target will become salient in the sparse part. This can be modeled as Equation 2:

$$\min_{B, T} \text{rank}(B) + \lambda \|T\|_0, s. t. D = B + T \quad (2)$$



where D is the raw image I , B is the low rank background image I_B , and T is the sparse foreground image I_T .

However, it is NP-hard to solve this problem since the rank function and l_0 -norm are both non-convex and discontinuous. Many researchers have used another relaxed form, as shown in Equation 3:

$$\min_{B,T} \|B\|_* + \lambda \|T\|_1, s.t. D = B + T \quad (3)$$

Here, the nuclear norm is a relaxation of the rank function, and the l_1 -norm is a relaxation of the l_0 -norm.

When we consider the effect of noises, the augmented Lagrange multipliers of this problem can be written as Equation 4:

$$L(B, T) = \|B\|_* + \lambda \|T\|_1 + \frac{\mu}{2} \|D - B - T\|_F^2 \quad (4)$$

Then we can solve it via some optimization algorithms such as the Accelerated Proximal Gradient (APG) (Lin et al., 2009). Especially, to deal with complex backgrounds better, researchers usually don't directly separate the raw image, but firstly construct the IPI or the tensor data, then decompose it. Taking IPI for an example, the flowchart of the algorithm is shown in Figure 3.

Firstly, put a window on the raw image and slid it at a given step to obtain a series of image patches. Then, these patches are flattened into vectors and formed as a new matrix \mathcal{D} , and decomposition will be applied on \mathcal{D} according to Equation 5 to get the low rank matrix \mathcal{B} and the sparse matrix \mathcal{T} .

$$L(\mathcal{B}, \mathcal{T}) = \|\mathcal{B}\|_* + \lambda \|\mathcal{T}\|_1 + \frac{\mu}{2} \|\mathcal{D} - \mathcal{B} - \mathcal{T}\|_F^2 \quad (5)$$

Finally, a median function will be used between some overlapped patches to recover the low rank background matrix B and sparse target matrix T .

Researchers usually tend to directly find small dim targets in the sparse matrix, however, when the background is very complex, due to the sparsity of edge and corner information, they will more likely be separated into the sparse matrix, overwhelming the target. In our work, we will turn to the low rank matrix and take it as the benchmark for global contrast information calculation. The

problem has become how to maintain as much edge/corner information as possible in the low rank matrix.

Let's analyze the reason for the data changes at complex edge and corner positions during the low rank conversion. Considering two simple cases and ignore the IPI operation, as shown in Figure 4: in (a), the raw image contains an edge and its rank is 5, if we forcefully reduce its rank to 3, two pixels will be changed and their gray value will be partially separated into the sparse image, see Figure 4b or Figure 4c.

Due to the self-similarity of pixel values in flat background areas, it can be reasonably inferred that these data changes should only occur at edges and corners, and within a small local area, as there are usually some constraints in the objective function during iterations. Therefore, we state that some local operations can be introduced on the global low rank background to maintain as much edge and corner information as possible, and apply a simple but effective local max dilation (LMD) operation on the separated low rank image.

The procedure for the global background separation in this paper is shown in Algorithm 1. The IPI model is adopted here.

Input: The input raw IR image I with resolution $M \times N$, patch size $p \times q$, the parameter λ using in APG, and the dilation radius r in LMD.

Output: The global background image B_{LMD} after local max dilation.

- 1: Slide the patch window on the raw image to obtain a series of image patches, then stretch them into vectors and form them as a new matrix D .
- 2: Decompose D into a low rank part B and a sparse part T using APG algorithm:
- 3: $B_0 = B_{-1} = 0$; $T_0 = T_{-1} = 0$; $a_0 = a_{-1} = 1$; $\mu_0 > 0$; $\mu_a > 0$; $\eta < 1$.
- 4: **while** not converged **do**
- 5: $Y_k^B = B_k + \frac{a_{k-1}}{a_k}(B_k - B_{k-1})$
- 6: $Y_k^T = T_k + \frac{a_{k-1}}{a_k}(T_k - T_{k-1})$
- 7: $G_k^B = Y_k^B - 12(Y_k^B + Y_k^T - D)$
- 8: $(U, S, V) = \text{svd}(G_k^B), B_{k+1} = US_{\frac{\mu_k}{2}}[S]V^t$, where svd means the singular value decomposition, S means a soft-thresholding operation, and V^t is the transpose of the matrix V .

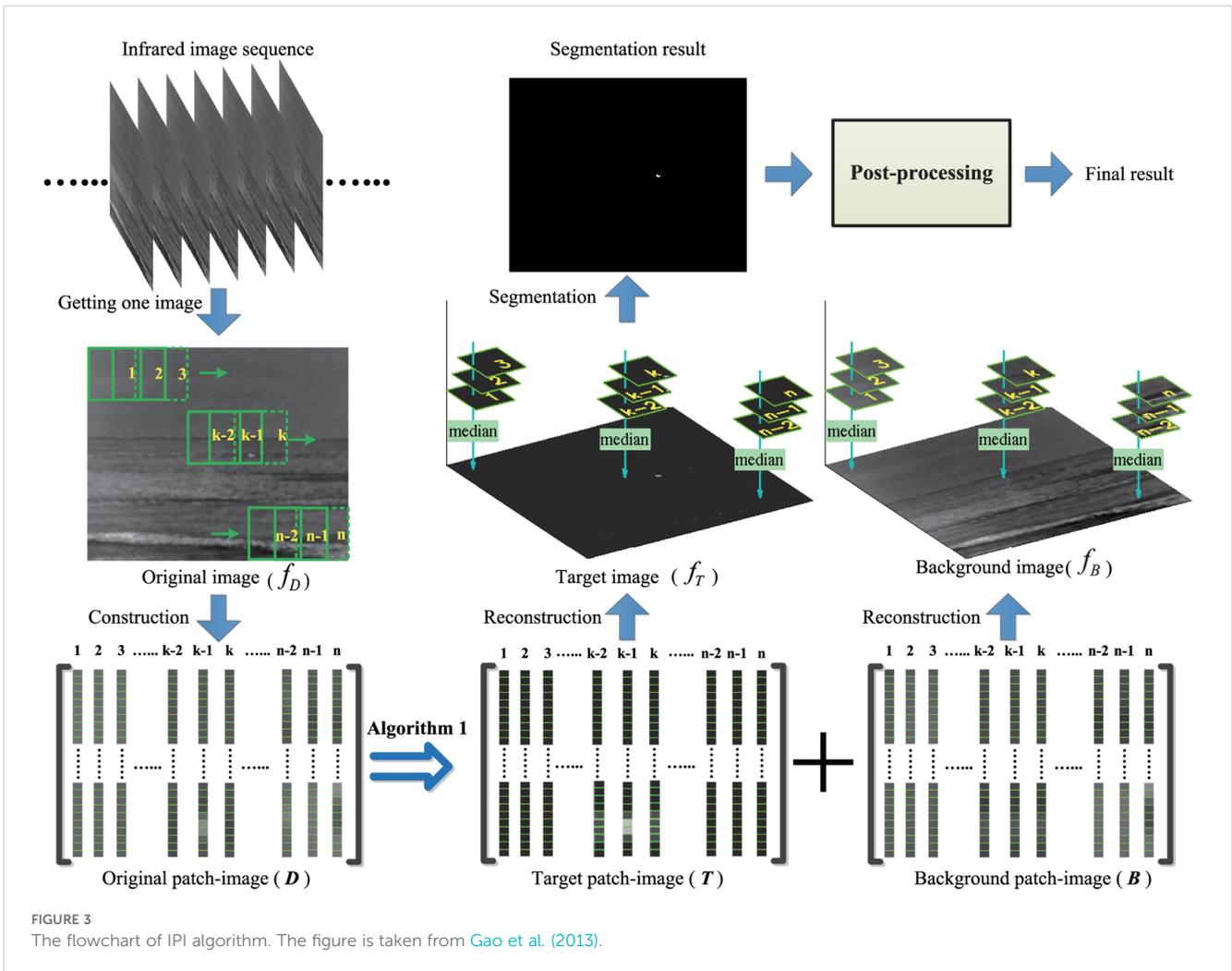


FIGURE 3 The flowchart of IPI algorithm. The figure is taken from Gao et al. (2013).

```

9:  $G_k^T = Y_k^T - 12(Y_k^B + Y_k^T - D)$ 
10:  $T_{k+1} = S_{\mu_k}[G_k^T]$ 
11:  $a_{k+1} = \frac{S_{\mu_k}[G_k^T]}{1 + \sqrt{4a_k^2 + 1}}, \mu_{k+1} = \max(\eta\mu_k, \mu_a)$ 
12:  $k = k + 1$ 
13: end while
14: Recover low rank global background image B from B:
 $B(x, y) = \text{median}(PB_i(x, y))$ ,  $PB_i \in B$  where  $PB_i$  is the  $i$ -th patch
containing position  $(x, y)$ , see Figure 3.
15: LMD operation for B:
6: for  $x=1:M$ 
17: for  $y=1:N$ 
18:  $B_{LMD}(x, y) = \max(B(x-r:x+r, y-r:y+r))$ 
19: end
20: end
21: Output  $B_{LMD}$ .
    
```

Algorithm 1. The global background separation.

In LMD, the dilation radius r is a key parameter. The larger the r is, the more the edge/corner information will be maintained. However, when a target approaches to high brightness

background and r is too large, the dilated high background will cover the target location, thus submerging the target during next step of contrast calculation. After a lot of experiments, we decided to set r to 1 in this paper to avoid this situation as much as possible.

2.1.2 Matched filtering

Random noises, especially PNHB, may also easily be separated into sparse foreground, and it is very hard to maintain them in the low rank background image by LMD operation since they usually emerge as single pixels. In this paper, before calculating the contrast information between raw image I and the benchmark B_{LMD} , a matched filtering operation will be applied on the raw image to suppress random noises first.

The theory of matched filter tells us that when the filter template is similar to the signal shape, the SNR of an image can be maximally improved (Moradi et al., 2016). Since true IR small dim targets usually has a Gaussian shape, a typical normalized Gaussian filtering template (Figure 5) will be utilized on the raw image to suppress random noises, as described in Equation 6:

$$I_{GS}(x, y) = \sum_{l=-1}^1 \sum_{k=-1}^1 GS(l, k) I(x + l, y + k) \tag{6}$$

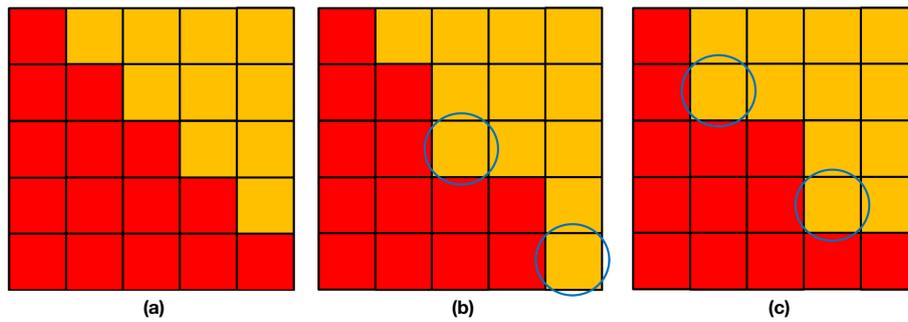


FIGURE 4

The low rank conversion at edges or corners. (a) A 5×5 image with rank of 5 containing a diagonal edge. (b) The image after conversion with rank of 3. (c) Another case with rank of 3.

where I is the raw image, (x, y) is the coordinate of each position in the raw image, GS is the Gaussian template in Figure 5, and I_{GS} is the filtering result.

We can also explain Equation 6 as a weighted gray sum in a small local area. A true target usually has an area larger than one pixel due to the point spread function of the optical system, so if there is a true target at the current location, the filtering result will be large. If there is a random noise with gray value equal to or slightly larger than the true target, the filtering result will be small, because a noise usually emerges as a single pixel.

2.1.3 Definition of RGCM

Inspired by our former work RLCM in the field of local contrast methods, the RGCM in this paper is defined as Equation 7:

$$\text{RGCM}(x, y) = \max \left\{ 0, \frac{I_{GS}(x, y)}{\max[\tau, B_{\text{LMD}}(x, y)]} I_{GS}(x, y) - I_{GS}(x, y) \right\} \quad (7)$$

The small value τ here is used to avoid division by zero, in this paper it is set to 5 for an 8-bit digital image. Meanwhile, there is a non-negative constraint to suppress clutters.

It also can be written as Equation 8:

$$\text{RGCM}(x, y) = \max[0, f(x, y)I_{GS}(x, y) - I_{GS}(x, y)] \quad (8)$$

Where f is defined as Equation 9:

$$f(x, y) = \frac{I_{GS}(x, y)}{\max[\tau, B_{\text{LMD}}(x, y)]} \quad (9)$$

It can be taken as an enhancement factor of the current pixel.

2.2 Small patch local contrast weighting function

In this paper, the DoG filter is selected as a weighting function for RGCM. DoG is a small patch local contrast method and can extract the local contrast information within a small area, thereby avoiding the problem of the targets being submerged by high

brightness backgrounds when approaching them. Similar to the original DoG method, we use two 5×5 templates (as shown in Figure 6) as the approximation of the Gaussian kernels, and define the weighting operation as Equation 10:

$$W = \max[I * (T_1 - T_2), 0] \quad (10)$$

Note that considering a desired target is usually hotter than the environment, a simple non-negative operation is utilized in the weighting function in this paper to suppress clutters better.

2.3 HCM calculation and discussions

Finally, the HCM is defined as Equation 11:

$$\text{HCM}(x, y) = \text{RGCM}(x, y)W(x, y) \quad (11)$$

It is obvious that both local and global contrast information are utilized in the proposed HCM method, that is why it is called as “hybrid” contrast method.

Algorithm 2 gives the main steps for HCM calculation.

Input: A raw input image I with resolution $M \times N$.

Output: The HCM matrix.

1: Decompose I according to Algorithm 1.

1: **for** $x=1:M$

2: **for** $y=1:N$

4: Calculate $I_{GS}(x, y)$ according to Equation 6.

5: Calculate $\text{RGCM}(x, y)$ according to Equation 7.

6: Weight $\text{RGCM}(x, y)$ according to Equation 10 and Equation 11 to get the $\text{HCM}(x, y)$.

8: **end**

9: **end**

10: Normalize HCM matrix to $0 \sim 1$ and output it.

Algorithm 2. HCM calculation.

It is necessary to discuss the different cases where (x, y) is different types of pixels:

There will be

$$B_{\text{LMD}}(x, y) \geq I_{\text{GS}}(x, y) \quad (24)$$

So, we can get that

$$f(x, y) \leq 1 \quad (25)$$

$$\text{RGCM}(x, y) \approx 0 \quad (26)$$

And

$$\text{HCM}(x, y) \approx 0 \quad (27)$$

Equations 22–27 means that the proposed method can effectively suppress complex background edges and corners.

d) If (x, y) is a random noise, the case will be similar to the true target. However, since the Gaussian filtering operation can suppress single pixel random noise to some extension, it can be easily deduced that the HCM of a noise pixel will be smaller than a true target's, even if its gray value is equal to or slightly larger than the target. Therefore, the HCM method can suppress random noise effectively.

2.4 Threshold operation

For each pixel of the raw IR image, calculate its HCM and form them as a new matrix. It is obvious that the true target will be the most salient in the HCM result, while other interferences such as edges, corners and noises are all inhibited. Therefore, in this paper the HCM result will be treated as the Saliency Map (SM), and an adaptive threshold operation will be used to extract the true target from it. The threshold value Th is defined as Equation 28

$$Th = \xi \cdot \max_{\text{SM}} + (1 - \xi) \text{mean}_{\text{SM}} \quad (28)$$

where \max_{SM} and mean_{SM} are the maximum and mean value of SM, respectively. ξ is a factor range $0 \sim 1$, according to our experiments, a ξ range from 0.7 to 0.95 is proper for most cases of single-target detection, but note that it's better to set ξ to a smaller value for multi-target detection cases, since different targets may have different saliency.

By applying the threshold Th on SM, the pixels larger than Th will be labeled as 1, otherwise labeled as 0. In the final binary image, output each connected area as a detected target (to eliminate clutters better, a dilation operation may be needed first).

Algorithm 3 summarizes the main steps for the threshold operation.

Input: The saliency map SM with resolution $M \times N$, and the parameter ξ .

Output: The target positions.

1: Calculate the threshold Th according to Equation 28.

2: **for** $x=1:M$

3: **for** $y=1:N$

4: **if** $\text{SM}(x, y) < Th$

5: $\text{SM}(x, y) = 0$

6: **else**

7: $\text{SM}(x, y) = 1$

8: **end**

9: **end**

10: **end**

11: Apply a dilation operation on SM to suppress clutters.

12: Output the connected areas in SM as target positions.

Algorithm 3. Threshold operation.

3 Details of the IR image data used in this paper

In this paper, 14 real IR sequences which contain different types of targets and backgrounds are used to verify the performance of the proposed algorithm. Figure 7 shows some samples of them, and Table 1 reports the details.

Besides, a single frame dataset is also used to test the detection performance of the proposed method, some samples are shown in Figure 8.

From Figures 7, 8 it can be seen that in the raw IR images, the targets are usually very small and dim, while the backgrounds are usually very complex. Besides, some images contain heavy noises.

4 Experiments and results

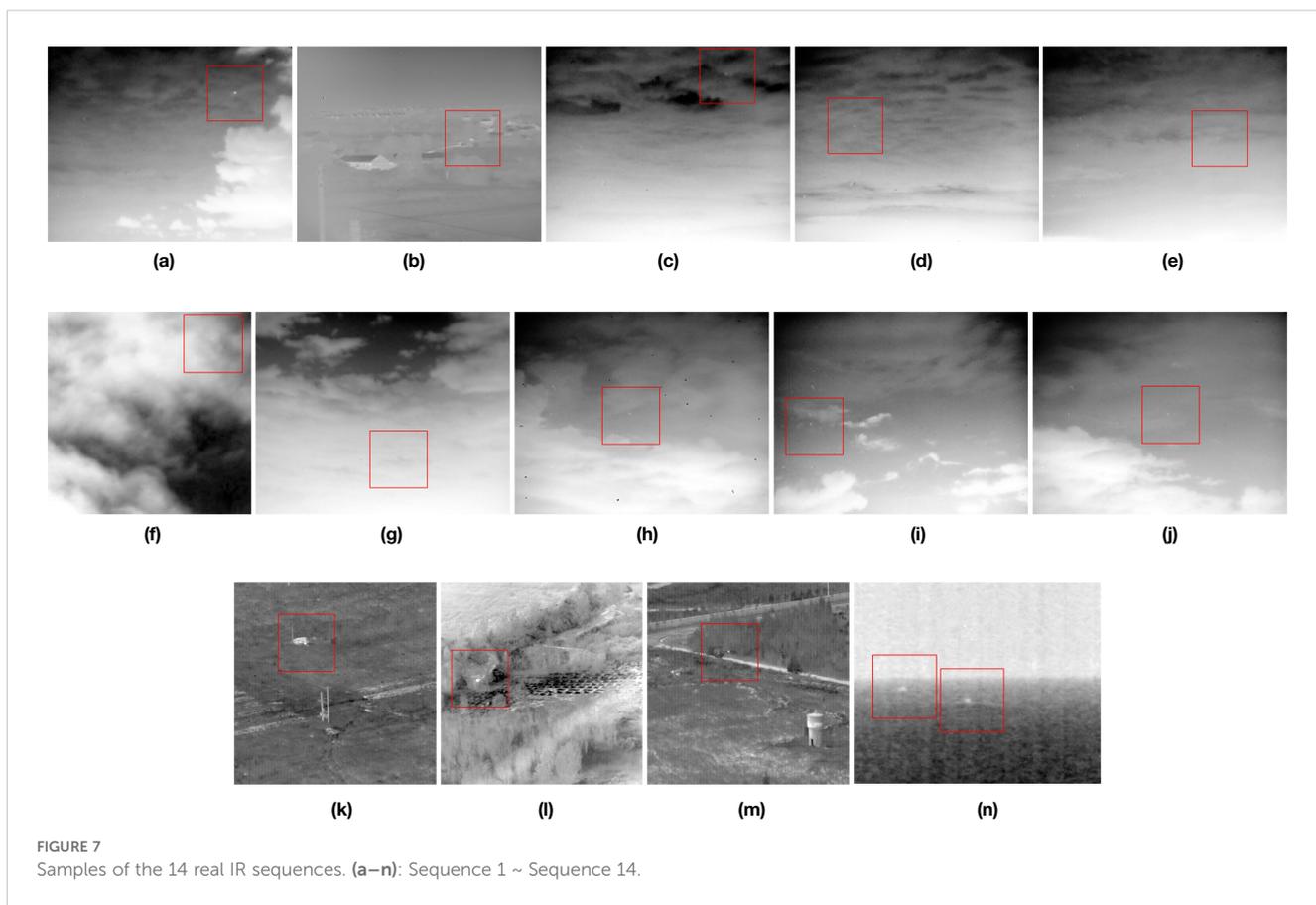
In this section, we will firstly give each processing step of the proposed method, and compare the performance of the proposed algorithm with some baseline algorithms. Then, the computational complexity and time consuming of the proposed method will be analyzed, and the robustness to noises of the proposed algorithm will be tested. Finally, some ablation experiments will be conducted to verify the effectiveness of some important modules of the proposed method. All the experiments are conducted on a PC with 8-GB random access memory and 3.1-GHz Intel i5 processor.

4.1 Processing results of the proposed algorithm

Firstly, the detection ability of the proposed algorithm are tested and the processing results of the proposed algorithm are given step by step in Figures 9, 10. The same samples with Figures 7, 8 are given here.

It can be seen from Figures 9, 10 that:

In the original IR image I , the targets are usually small and dim, while backgrounds are usually complex, they may have high brightness and complex edges and corners. Meanwhile, there are many random noises (including PNHB) in some sequences, too.



The Gaussian filter in the first step can effectively suppress random noises and improve image quality to a certain extent.

After sparse and low rank decomposition, the global background image B separated from the original image mainly contains background information. However, it lost some important information of complex edges/corners, which makes it not suitable to be directly considered as the benchmark for global contrast calculation in next steps.

After LMD, the dilated global background image B_{LMD} can recover as much edges/corners information as possible, so it is more suitable to be used as the benchmark than B .

After RGCM calculation, real targets become very salient. However, there are still a few clutters in some images with complex backgrounds.

After the weighting operations by the DoG filtering result, residual clutters are suppressed further and real targets become the most salient in the SM.

Finally, after the threshold operation, all the real targets are output successfully, and only one false alarm emerges in Sequence 7 (it is a broken cloud which has a similar pattern to the real target, in our future work, we will utilize some time domain methods to eliminate it). Therefore, the effectiveness of the proposed method is proved.

4.2 Comparisons with other algorithms

Nine existing algorithms are chosen as baselines for comparisons to verify the advantages of the proposed method, including:

Seven local contrast algorithms, such as DoG (Wang et al., 2012), ILCM (Han et al., 2014), MPCM (Wei et al., 2016), RLCM (Han et al., 2018), Weighted Local Difference Measure (WLDM) (Deng et al., 2016), Multi-Directional Two-Dimensional Least Mean Square (MDTDLMS) (Han et al., 2019), and Enhanced Closest Mean Background Estimation (ECMBE) (Han et al., 2021).

One global decomposition algorithm, i.e., IPI (Gao et al., 2013).

One deep learning algorithm, i.e., RPCANet (Wu et al., 2024).

Here is a summary of each baseline method:

- DoG is a traditional small patch local contrast method.
- ILCM is a large patch local contrast method, it takes the ratio value between a current pixel and its surrounding benchmark as contrast information.
- MPCM is a large patch local contrast method, but it performs multiscale calculation to extract the target better.
- RLCM is a multiscale local contrast method too, and both ratio and difference operations are utilized in it to enhance true target and suppress background simultaneously.

TABLE 1 Features of the 14 real IR sequences.

Sequence	Frames	Image Resolution	Target ID	Target Size	Target Details	Background Details
1	330	320×256	Only 1	About 4×5	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Some bad pixels. Almost unchanged.
2	400	320×256	Only 1	Varies from 2×2 to 3×5	<ul style="list-style-type: none"> Car target. A long imaging distance. Located on road. Keeping motion. 	<ul style="list-style-type: none"> Building-Tree background. Heavy clutters including bright areas. Some bad pixels. Almost unchanged.
3	330	320×256	Only 1	About 3×5	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Many bad pixels. Almost unchanged.
4	300	320×256	Only 1	About 3×5	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Some bad pixels. Almost unchanged.
5	200	320×256	Only 1	About 3×3	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Some bad pixels. Almost unchanged.
6	267	250×250	Only 1	Varies from 2×3 to 3×5	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping motion. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Almost unchanged.
7	280	320×256	Only 1	About 3×4	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Many broken clouds. Some bad pixels. Almost unchanged.
8	260	320×256	Only 1	About 3×3	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Many bad pixels. Change slowly.
9	250	320×256	Only 1	Varies from 2×2 to 3×3	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Some bad pixels. Change slowly.
10	370	320×256	Only 1	Varies from 2×2 to 3×3	<ul style="list-style-type: none"> Plane target. A long imaging distance. Located in cloudy sky. Keeping little motion. Very small and dim. 	<ul style="list-style-type: none"> Sky-Cloud background. Heavy clouds including bright areas. Complex textures. Some bad pixels. Change slowly.

(Continued)

TABLE 1 Continued

Sequence	Frames	Image Resolution	Target ID	Target Size	Target Details	Background Details
11	330	256×256	Only 1	Varies from 2×3 to 3×3	<ul style="list-style-type: none"> • Unmanned aerial vehicle target. • Located in sky. • Moving fast. • Near to high brightness background in some frames. 	<ul style="list-style-type: none"> • Ground background. • Heavy clutters including bright areas. • Complex edges and corners. • Changed fast.
12	470	256×256	Only 1	Varies from 3×3 to 3×4	<ul style="list-style-type: none"> • Unmanned aerial vehicle target. • Located in sky. • Moving fast. 	<ul style="list-style-type: none"> • Ground-Tree background. • Heavy clutters including bright areas. • Complex edges and corners. • Changed fast.
13	230	256×256	Only 1	Varies from 3×3 to 3×5	<ul style="list-style-type: none"> • Unmanned aerial vehicle target. • Located in sky. • Moving fast. 	<ul style="list-style-type: none"> • Ground-Tree background. • Heavy clutters including bright areas. • Complex edges and corners. • Changed fast.
14	300	280×228	Target 1	About 5×6	<ul style="list-style-type: none"> • Ship target. • A long imaging distance. • Located in homogeneous sea. • Two targets, one is moving and the other is stationary. 	<ul style="list-style-type: none"> • Sea-Sky background. • Heavy wave clutters. • Heavy noises. • Almost unchanged.
			Target 2	About 5×5		

- e. WLD introduces the local entropy as a weighting function for the local contrast information.
- f. MDTDLMS utilizes a background estimation method to get the benchmark for contrast information calculation, however, its benchmark is still obtained by some local operations.
- g. ECMBE combines local contrast method with background estimation too, and it proposes a new background estimation principle named closest mean, therefore the problem of target submergence caused by the neighboring high brightness background can be alleviated.
- h. IPI is a sparse and low rank decomposition algorithm, but it focuses on the sparse foreground image and tries to directly search target in it.
- i. RPCANet is a newly proposed deep learning algorithm, it unfolds the traditional iterations in RPCA algorithm with deep networks to achieve sparse and low rank decomposition.

The key parameters of each algorithm are listed in Table 2.

Firstly, two objective indicators named SCR Gain (SCRG) and Background Suppression Factor (BSF) are used to describe different performances of different algorithms. SCRG, which is defined as Equation 29, can effectively describe the target enhancement ability of an algorithm. BSF, which is defined as Equation 30, can describe the background suppression ability of an algorithm.

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}} \quad (29)$$

$$\text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}} \quad (30)$$

where SCR_{in} and SCR_{out} are the SCR (defined as Equation 31) of the raw image and SM respectively, σ_{in} and σ_{out} are the standard deviation of the raw image and SM respectively.

$$\text{SCR} = \frac{|I_t - I_{\text{nb}}|}{\sigma} \quad (31)$$

where I_t is the maximal gray of the target center. I_{nb} is the average gray of the neighboring background around the target

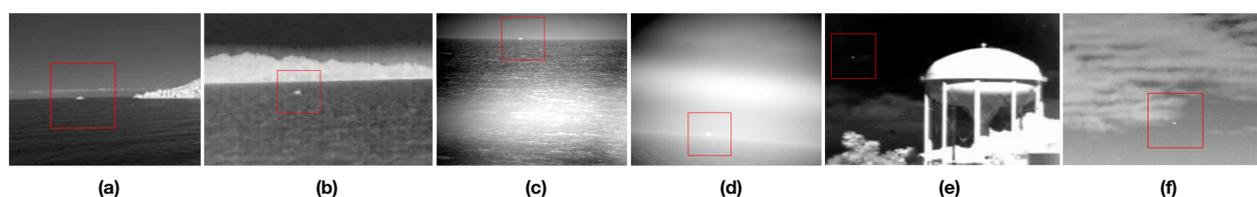


FIGURE 8 Six samples of the single frame dataset. (a-f): sample 1 ~sample 6.

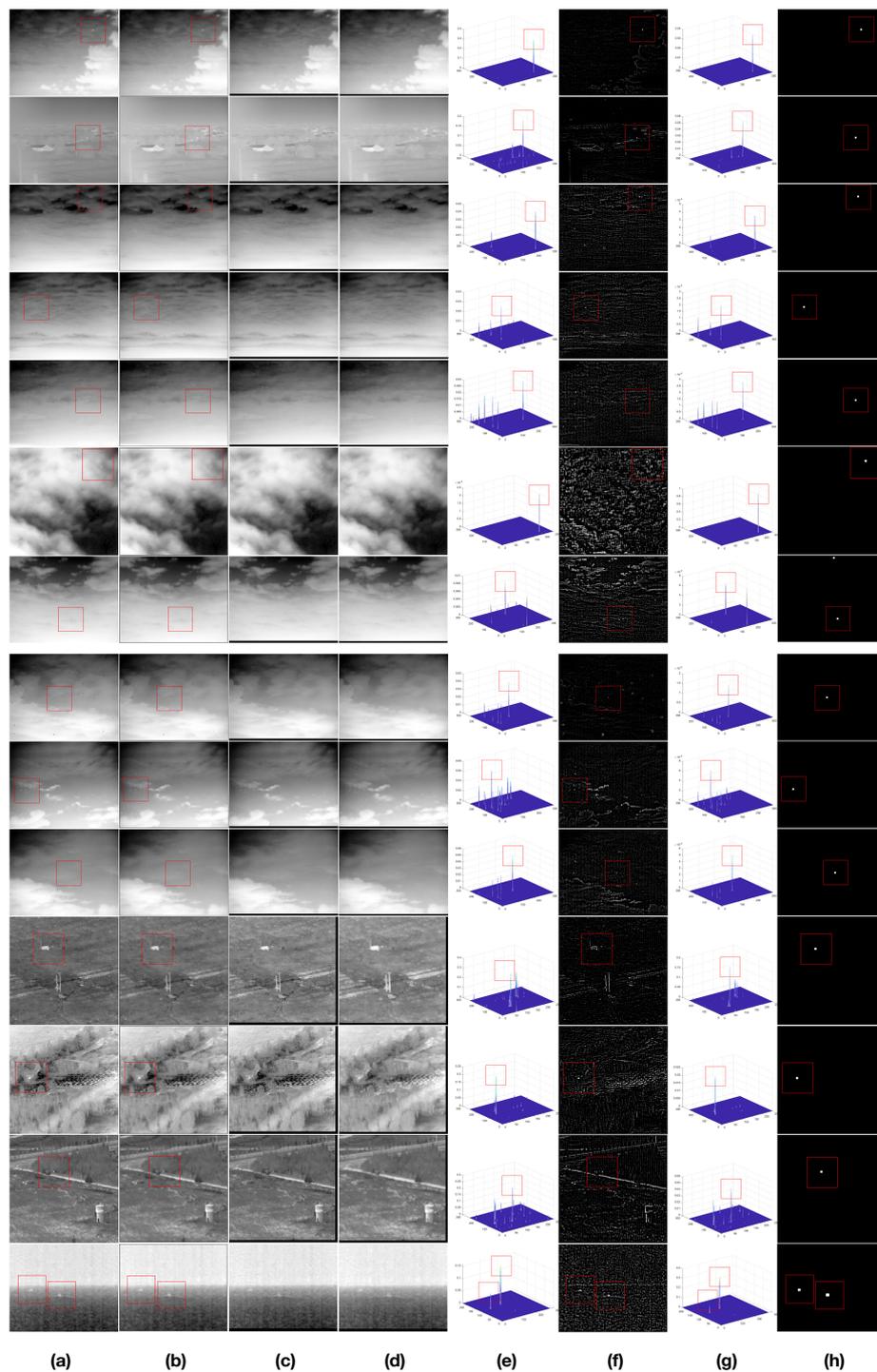


FIGURE 9

The processing results of the proposed algorithm for the samples of the 14 sequences, from top to bottom: Sequence 1~ Sequence 14. (a) Raw IR images. (b) Images after Gaussian filtering. (c) The separated global background B. (d) The global background B_{LMD} after LMD. (e) The RGCM results. (f) The DoG weighting function. (g) The HCM results after weighting. (h) The final detection results after threshold operation.

center, in this paper it is set to the area between 15×15 and 9×9 around the target center. σ is the stand deviation of the image.

The results of SCRG and BSF are shown in Tables 3, 4, respectively. Please note that as a deep learning method, the output of RPCANet is the target probability of each pixel, so we are unable to calculate its SCRG and BSF.

It can be seen from Tables 3, 4 that compared to the baselines, the proposed algorithm can achieve the highest SCRG and BSF in most cases.

Then, to intuitively show the detection performance of different algorithms, Figures 11, 12 give the saliency map and the detection results of each algorithm for the samples of the 14 sequences.

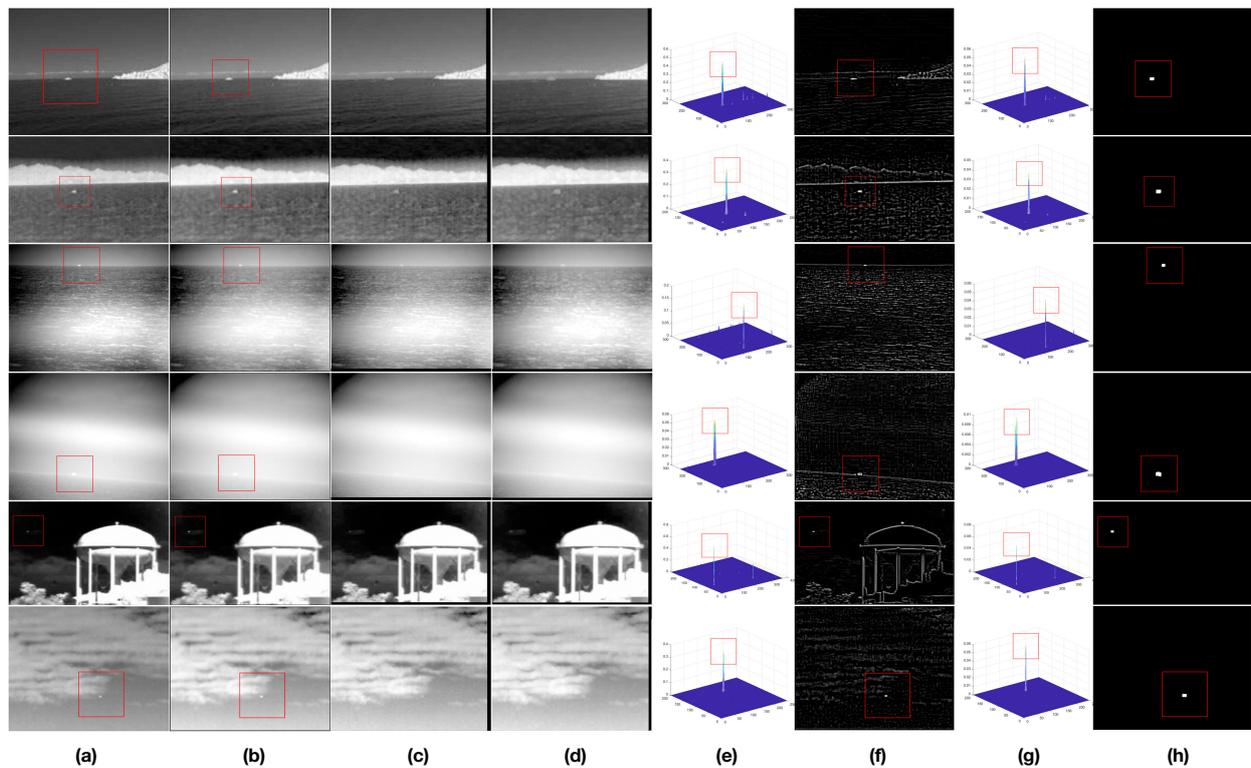


FIGURE 10

The processing results of the proposed algorithm for the samples of the single frame dataset. (a) Raw IR images. (b) Images after Gaussian filtering. (c) The separated global background B . (d) The global background B_{LMD} after LMD. (e) The RGCM results. (f) The DoG weighting function. (g) The HCM results after weighting. (h) The final detection results after threshold operation.

Besides, the Receiver Operating Characteristic (ROC) curves for each whole sequence are utilized to compare the detection performance of different algorithms, and the results are shown in Figure 13. Here, the False Positive Rate (FPR) and the True Positive Rate (TPR) are defined as Equation 32 and Equation 33.

$$FPR = \frac{\text{number of detected false targets}}{\text{total number of pixels in the whole image}} \quad (32)$$

$$TPR = \frac{\text{number of detected true targets}}{\text{total number of real targets}} \quad (33)$$

It can be seen from Figures 11-13 that:

- DoG can only extract true targets in Sequence 1, Sequence 3 and Sequence 14 in Figure 12, and its ROC performance is usually the worst in Figure 13, too.
- ILCM can extract true targets in Sequence 1, Sequence 7, Sequence 10, Sequence 12 and Sequence 14, its performance is slightly better than DoG. However, it is not satisfied in many other sequences.
- MPCM can output true targets in Sequence 1, Sequence 2, Sequence 12 and Sequence 14. However, when the background is complex, many interference are enhanced and output too, for example, in Sequence 4, Sequence 6, Sequence 7 and Sequence 9, etc. Especially, in Figure 13, its performance is worse than many other algorithms.
- RLCM can achieve a better detection performance in some sequences, for example, in Sequence 4, Sequence 8 and Sequence 9, etc. However, when the background is very complex and the target is very dim, it will fail, for example, in Sequence 5, Sequence 6, Sequence 11 and Sequence 13, etc.
- WLDM utilizes the local entropy as the weighting function for local contrast information, however, when the target is dim and the background is very complex, the target will be submerged by clutters, for example, in Sequence 3 ~ Sequence 10, etc.
- MDTDLMS utilizes the TDLMS background estimation method to get the benchmark for contrast information calculation, so it can achieve good performance in some cases, such as in Sequence 3, Sequence 5, Sequence 12 and Sequence 13, etc. However, its benchmark is obtained locally, so its performance is still not good in some cases, especially when the target approaching some high brightness background, for example, in Sequence 11.
- ECMBE improved the principle of benchmark selection, so it can achieve good detection performance even if the target is not local salient, for example, in Sequence 11. However,

TABLE 2 The parameter values used in the baseline algorithms.

	Parameter values
DoG	A 5×5 binomial kernel as author recommended. $C_{Tn}=0.15$ for single-target situation and 0.01 for multi-target situation.
ILCM	A same 5×5 binomial kernel with DoG, and the cell size is 8×8.
MPCM	Three scales with cell size 3×3, 5×5 and 7×7 are used.
RLCM	The cell size is 9×9. Three scales are used, and (K_1, K_2) is set to (2, 4), (5, 9) and (9, 16), respectively.
WLDM	Entropy window size is 5×5. Four scales are used for LDM calculation, and the cell size is 3×3, 5×5, 7×7 and 9×9, respectively.
MDTDLMS	The inner window is 7×7, the outer window is 11×11, and μ is set to 10^{-7} for 8-bit images.
ECMBE	The central layer is 3×3, the isolating layer is 7×7, and the surrounding layer is 7×7.
IPI	The patch size is 50×50, the step is 10, λ is $1/\sqrt{\max(m,n)}$.
RPCANet	The trained parameters of the network are downloaded from the authors' github.

TABLE 3 The SCRG of different algorithms in the 14 sequences.

Sequence	Target	DoG	ILCM	MPCM	RLCM	WLDM	MDTDLMS	ECMBE	IPI	RPCANet	Proposed
1	1	6.316	28.763	4.878	12.686	90.853	45.971	70.676	76.028	–	97.014
2	1	0.548	1.133	2.087	1.290	12.639	4.869	10.119	11.465	–	22.213
3	1	8.935	17.292	3.690	16.557	40.904	45.075	128.828	82.053	–	300.988
4	1	5.676	22.872	1.084	8.076	46.389	32.411	127.327	87.506	–	351.813
5	1	10.645	10.011	2.303	9.952	98.753	74.875	121.774	147.741	–	333.891
6	1	4.890	6.007	1.260	5.120	7.000	0.422	75.084	161.634	–	591.153
7	1	23.462	57.587	7.123	8.088	45.541	34.230	98.213	240.984	–	1.016E3
8	1	4.722	1.779	0.628	9.397	6.324	49.381	142.654	25.871	–	355.369
9	1	4.120	13.185	1.085	7.167	42.672	23.193	59.347	79.010	–	202.156
10	1	5.644	36.641	0.353	6.294	58.622	40.036	92.332	107.412	–	244.940
11	1	0.643	0.638	1.486	0.494	3.820	1.563	2.208	3.843	–	11.220
12	1	0.960	3.099	5.310	4.391	4.671	11.277	17.274	14.610	–	42.769
13	1	0.574	0.426	0.518	0.526	3.852	3.435	4.040	9.390	–	19.922
14	1	10.799	14.609	11.128	6.127	34.068	35.088	48.003	60.438	–	41.034
	2	6.627	11.963	9.203	7.439	23.238	31.017	47.937	66.447	–	92.900

The bold values are the largest values.

since its benchmark is still obtained locally, its performance is not very good in some cases, for example, in Sequence 5 ~ Sequence 7, and Sequence 9 ~ Sequence 11, etc.

- h. As a global decomposition method, IPI can achieve good performance even when the target is not local salient, for example, in Sequence 11. However, it focus on sparse part and is sensitive to complex edge/corner information. If the target is weak and the background is complex, it will fail, for example, in Sequence 2 ~ Sequence 7, Sequence 10 ~ Sequence 13, etc.
- i. RPCANet, as a deep learning method, can achieve good performance when the data distribution is the same with the training samples, for example, in Sequence 1. However,

when the data distribution is different, its performance will decrease significantly.

Compared to the existing methods, the performance of the proposed HCM algorithm is always in the forefront in all of the 14 sequences. Especially, when the clutter is heavy the target is not local salient, it can still achieve a good detection performance.

4.3 Comparisons of computational complexity and time consuming

In this section, the computational complexity and time consuming for different algorithms is analyzed. For simplicity, suppose the raw image has a resolution of $X \times Y$, and the scale of

TABLE 4 The BSF of different algorithms in the 14 sequences.

Sequence	DoG	ILCM	MPCM	RLCM	WLDM	MDTDLMS	ECMBE	IPI	RPCANet	Proposed
1	3.617	27.820	0.011	7.181	11.040	3.785E5	5.808E5	1.037E6	–	2.093E7
2	0.430	6.092	0.145	1.752	18.794	3.346E3	6.926E3	1.585E4	–	2.776E5
3	5.731	55.319	0.019	15.708	24.602	1.349E5	3.815E5	7.604E5	–	5.435E7
4	3.301	54.694	0.020	11.892	15.932	8.391E4	2.966E5	6.207E5	–	3.102E8
5	7.045	95.262	0.030	15.683	41.273	1.370E5	2.314E5	5.652E5	–	2.073E8
6	4.834	116.547	0.190	6.246	656.763	1.131E3	2.628E3	7.414E3	–	1.680E4
7	4.232	89.609	0.012	12.175	22.301	1.287E5	2.755E5	7.490E5	–	9.332E8
8	3.428	3.975	0.019	11.142	1.898	1.485E5	3.620E5	4.621E5	–	5.540E8
9	2.473	46.555	0.011	6.601	8.589	1.913E5	2.673E5	4.731E5	–	7.440E7
10	4.179	69.465	0.011	8.364	21.241	1.311E5	3.246E5	6.378E5	–	1.298E8
11	0.175	3.490	0.215	0.850	24.653	345.695	483.090	625.085	–	6.940E3
12	0.351	8.474	0.155	2.816	32.035	1.856E3	2.827E3	3.667E3	–	2.877E5
13	0.229	4.593	0.159	1.087	23.780	387.889	539.851	1.528E3	–	4.043E4
14	3.734	63.267	1.464	6.264	1.792E3	924.493	1.415E3	2.179E3	–	8.282E3

The bold values are the largest values.

the patch window or cell is $(2L+1)^2$. For multi-scale local algorithms, such as MPCM, RLCM and WLDM, etc., denote S as the scale number, L_i ($i=1, 2, \dots, S$) is the L of the i_{th} scale.

For DoG, there will be $(2L+1)^2$ multiplications and $(2L+1)^2$ additions for each pixel during the convolution operation, so the computational complexity will be $O(L^2XY)$.

For ILCM, since it uses a DoG filter as preprocessing, and the latter subblock-stage processing consumes less calculations, its computational complexity will be $O(L^2XY)$.

For MPCM, for each scale, the average operation will cost $(2L_i+1)^2$ additions for each pixel, so for total S scales its computational complexity will be $O(SL_s^2XY)$.

For RLCM, for each scale, the sort operation within a cell will cost $(2L_i+1)^2 \log(2L_i+1)^2$ calculations, so for total S scales its computational complexity will be $O[SL_s^2 \log(L_s^2)XY]$.

For WLDM, for each scale, the average operation will cost $(2L_i+1)^2$ additions for each pixel, and the entropy calculation will need a sort operation within a cell first, which will cost $(2L_i+1)^2 \log(2L_i+1)^2$ calculations, so for total S scales its computational complexity will be $O[SL_s^2 \log(L_s^2)XY]$.

For MDTDLMS, its computational complexity is reported as $O(L^2XY)$ in the original paper.

For ECMBE, its computational complexity is reported as $O(LXY)$ in the original paper.

For IPI, the computational complexity is reported as $O(Nkmn \log(mn) + rc(p+1))$ in the original paper. Here, m is the number of pixels of the patch window, i.e., $(2L+1)^2$ in this paper. n is the number of patches, k is the number of nonzero singular value (rank) of G_k^T , N is the iteration number, p is the overlapping pixel number during the transformation from the target/background patch image to the reconstruction image. r and c are the row and column numbers of the original image, i.e., X and Y in this paper,

respectively. Therefore, its computational complexity can be rewritten as $O(NkL^2n \log(L^2n) + XY(p+1))$ here.

For RPCANet, since it is a deep learning method, its computational complexity is not given here.

The proposed algorithm has four main steps: global background separation, Gaussian matched filtering, RGCM calculation and DoG weighting operation. For global background separation, it has the same computational complexity as IPI algorithm; for the Gaussian matched filtering, it has 9 multiplications, 1 addition at each pixel, totally $10XY$ operations for the whole image; for RGCM calculation, there will be 8 comparisons for LMD of the background benchmark, 1 division, 1 subtraction, and 1 multiplication for relative GCM calculation, totally $11XY$ operations for the whole image; for the DoG weighting operation, there will be 25 multiplications and 1 addition for each pixel, totally $26XY$ operations for the whole image. Therefore, its computational complexity will be $O(NkL^2n \log(L^2n) + XY(p+47))$.

Table 5 summarizes the comparisons of computational complexity for different algorithms. The average time consuming (in seconds) of different algorithms for one frame is listed in Table 6. Please note that although some existing algorithms can achieve less computational complexity and average time consuming, their performances are too bad. The proposed algorithm, although didn't show advantages in computational complexity and time consuming, can achieve better detection performance.

4.4 The robustness to noises

To test the robustness to noises of the proposed algorithm, we select one sequence (Sequence 11) and add different levels of noises into it, then draw the ROC curves, see Figure 14. It can be seen that

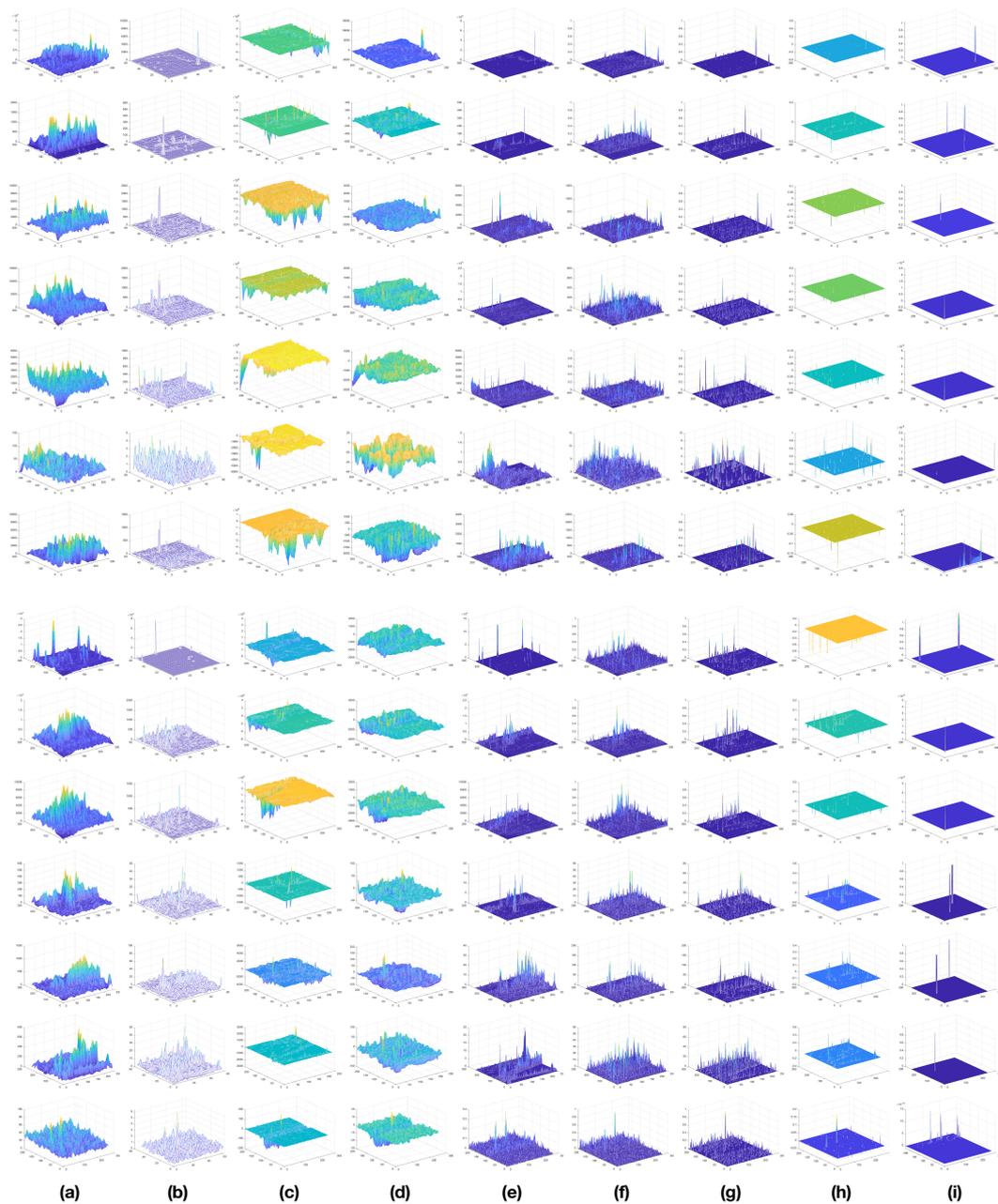


FIGURE 11

The saliency maps of the 14 sequences using different algorithms. (a) DoG. (b) ILCM. (c) MPCM. (d) RLCM. (e) WLDM. (f) MDTLMS. (g) ECMBE. (h) IPI. (i) RPCANet.

after the add of noise, the detection performance of the proposed algorithm doesn't change obviously.

4.5 The ablation experiments

To verify the effectiveness of some important modules, the ablation experiments are conducted at last. In the proposed algorithm, three modules are important: the Gaussian matched filter, the LMD for the low rank background, and the DoG

weighting function. All of them are tested and the results are given in Table 7.

From Table 7 we can see that the proposed algorithm with all of the three modules can achieve the best or the second best SCR and BSF in most cases, which proves the effectiveness of these modules for improving detection performance. It is worth noting that in some sequences, the algorithm that do not perform Gaussian filtering operation can achieve a larger SCR. This is because in these sequences the target is extremely small and the Gaussian filtering may smooth it to some extent. However, we still think the Gaussian

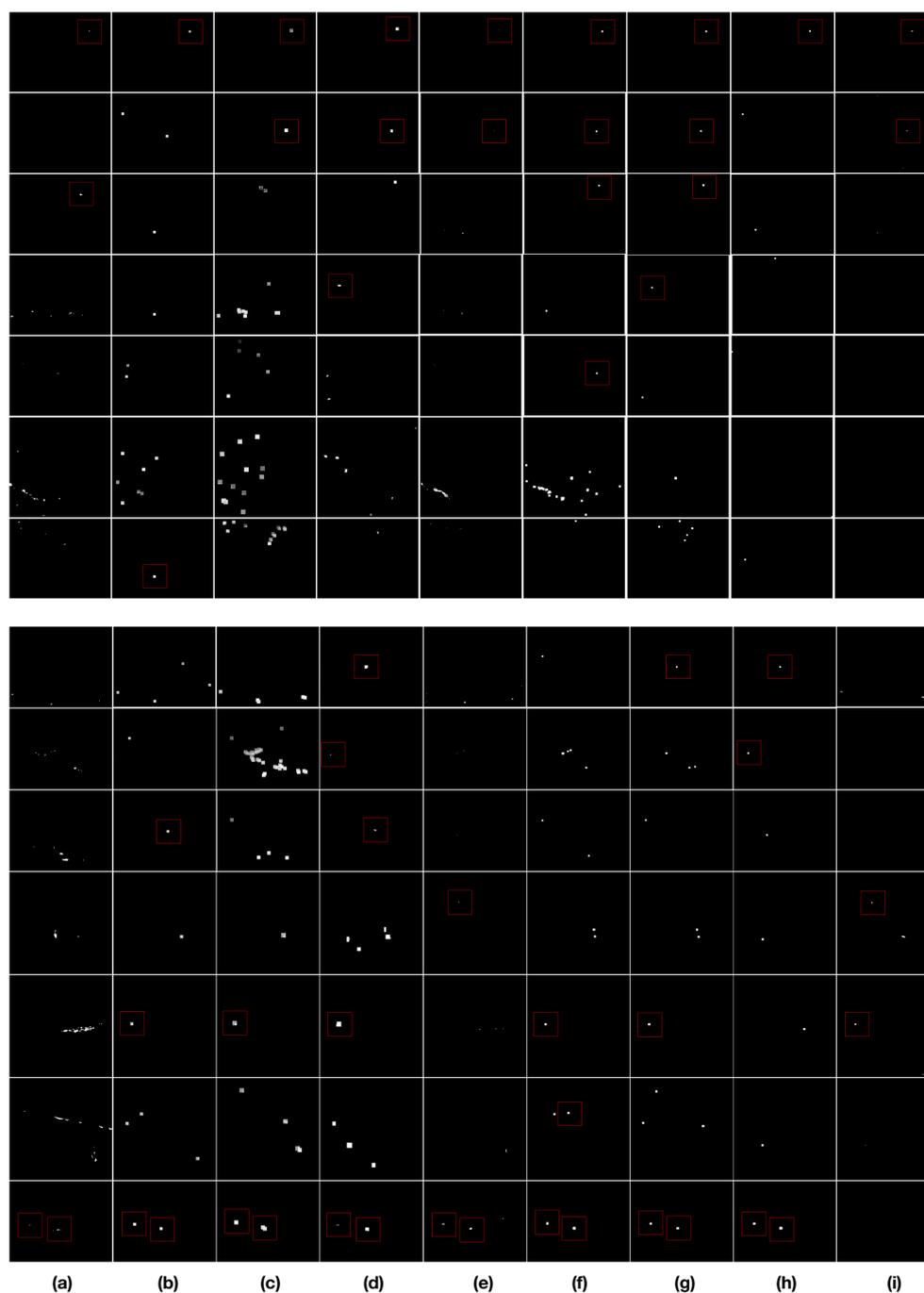


FIGURE 12

The detection results of the 14 sequences using different algorithms. (a) DoG. (b) ILCM. (c) MPCM. (d) RLCM. (e) WLDM. (f) MDTDLMS. (g) ECMBE. (h) IPI. (i) RPCANet.

filter is necessary since it can smooth backgrounds and noises too, that's why the algorithm without Gaussian filter has a smaller BSF.

5 Conclusions

In this paper, a new contrast method framework named hybrid contrast measure (HCM) is proposed for IR small target detection.

It consists of two types of contrast information: the global contrast, and the local contrast. In the global contrast calculation, it firstly obtains benchmark via a global sparse and low rank decomposition, so that it can handle the situation when target approaches to a high brightness background and becomes not local salient. Especially, a simple LMD operation is applied on the global low rank background benchmark to recover as much edge/corner information as possible. Then, the relative global contrast

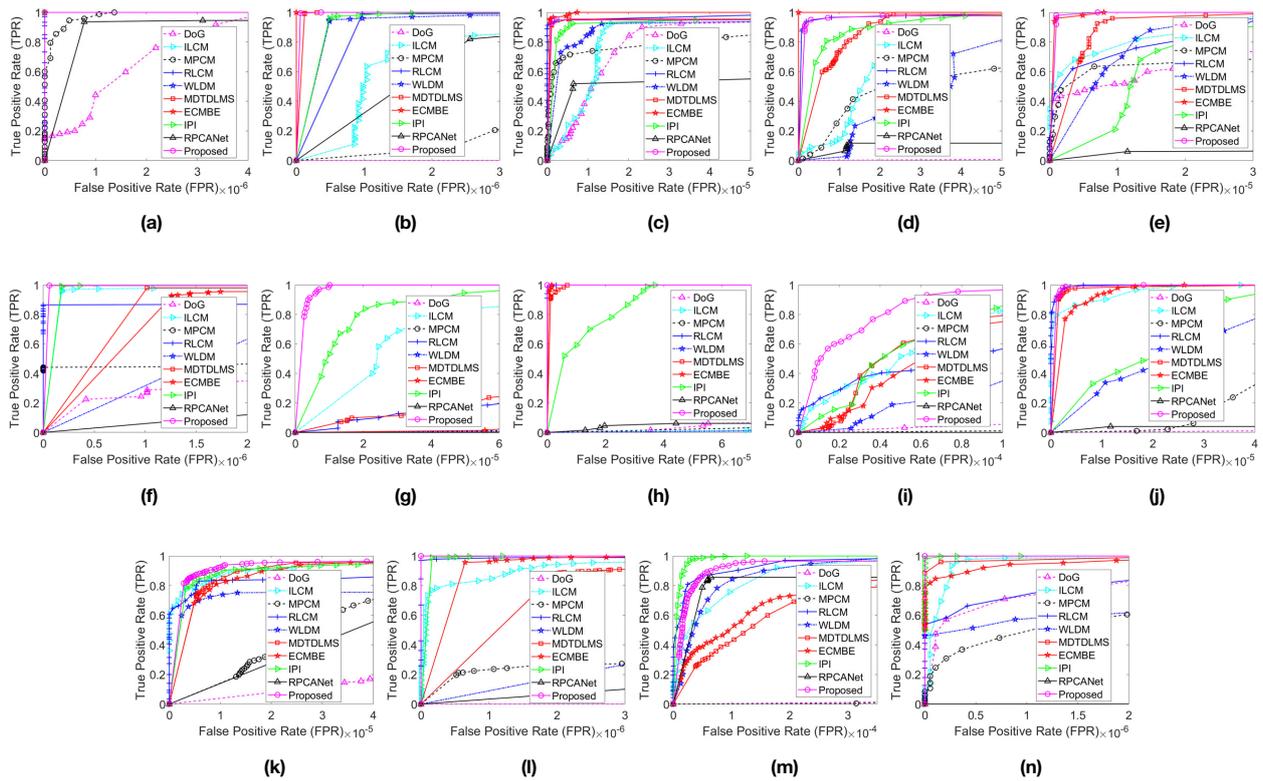


FIGURE 13 The ROC curves of different algorithms in different sequences. (a–n): Sequence 1 ~ Sequence 14.

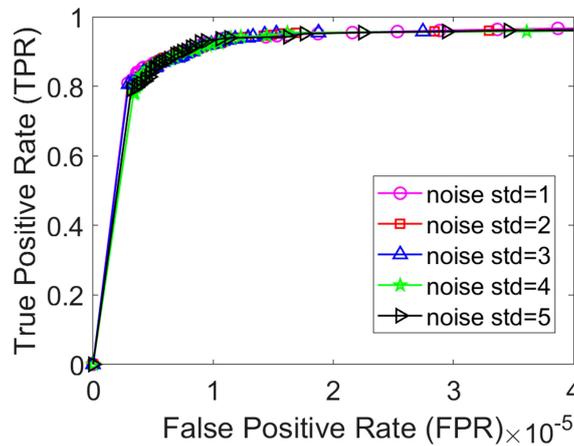


FIGURE 14 The detection performance of the proposed algorithm in Sequence 11 after different levels of noises are added.

TABLE 5 Computational complexity of different algorithms.

Algorithm	DoG	ILCM	MPCM	RLCM	WLDLMS
Computational Complexity	$O(L^2XY)$	$O(L^2XY)$	$O(SL_s^2XY)$	$O[SL_s^2 \log(L_s^2)XY]$	$O[SL_s^2 \log(L_s^2)XY]$
Algorithm	MDTDLMS	ECMBE	IPI	RPCANet	Proposed
Computational Complexity	$O(SL_s^2 \log(L_s^2)XY)$	$O(LXY)$	$O(NkL^2n \log(L^2n) + XY(p+1))$	-	$O(NkL^2n \log(L^2n) + XY(p+23))$

TABLE 6 Average time consuming of different algorithms for 14 sequences (seconds per frame).

Sequence	DoG	ILCM	MPCM	RLCM	WLDM	MDTDLMS	ECMBE	IPI	RPCANet	Proposed
1	0.113	0.035	0.918	1.202	1.450	3.420	0.203	8.544	0.536	8.695
2	0.074	0.023	0.737	1.355	2.108	10.555	0.507	15.651	0.531	13.750
3	0.190	0.023	0.612	1.254	1.322	3.602	0.177	12.108	0.563	5.380
4	0.081	0.031	1.115	1.198	2.012	3.027	0.336	4.905	0.571	5.158
5	0.064	0.028	0.934	1.331	1.766	3.030	0.276	6.288	0.521	6.213
6	0.057	0.025	0.896	1.467	1.438	2.683	0.253	4.566	0.386	4.680
7	0.080	0.033	0.882	1.054	1.058	3.037	0.189	6.983	0.540	7.255
8	0.110	0.032	0.956	0.967	1.115	2.788	0.301	7.082	0.570	7.196
9	0.102	0.019	1.774	0.884	1.082	3.023	0.244	5.778	0.520	5.891
10	0.055	0.022	0.669	1.348	1.153	3.019	0.179	7.150	0.571	7.249
11	0.062	0.021	0.892	1.043	1.133	5.829	0.185	3.314	0.389	3.446
12	0.054	0.024	1.003	1.103	1.158	3.648	0.165	3.143	0.460	3.378
13	0.066	0.032	0.925	1.130	1.093	4.595	0.203	3.165	0.369	3.301
14	0.053	0.018	0.476	0.628	1.056	2.734	0.136	2.703	0.545	2.273

TABLE 7 Ablation experiments.

Gaussian filter				✓		✓		✓	
LMD		✓				✓		✓	
Weighting operation		✓		✓				✓	
Sequence	Target	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
1	1	133.928	9.864E6	97.444	1.884E7	76.409	1.649E7	97.014	2.093E7
2	1	29.340	8.559E4	22.002	2.488E5	15.609	1.951E5	22.213	2.776E5
3	1	306.460	3.028E7	298.755	4.695E7	236.877	4.277E7	300.988	5.435E7
4	1	323.895	6.322E7	162.517	1.436E8	326.720	2.236E8	351.813	3.102E8
5	1	275.106	6.647E7	288.026	1.350E8	262.833	1.580E8	333.891	2.073E8
6	1	450.003	1.538E4	144.636	1.598E4	591.153	1.680E4	591.153	1.680E4
7	1	1.031E3	2.428E8	714.261	3.018E8	1.031E3	7.125E8	1.016E3	9.332E8
8	1	366.159	1.325E8	300.761	4.057E8	277.094	1.548E8	355.369	5.540E8
9	1	232.863	2.583E7	140.887	5.026E7	119.682	4.407E7	202.156	7.440E7
10	1	268.457	3.977E7	215.553	1.092E8	190.242	1.008E8	244.940	1.298E8
11	1	7.912	4.052E3	10.604	4.886E3	6.599	4.091E3	11.220	6.940E3
12	1	21.766	1.369E5	37.281	2.137E5	35.896	2.407E5	42.769	2.877E5
13	1	21.040	1.097E4	15.731	2.729E4	16.862	3.323E4	19.922	4.043E4
14	1	48.215	7.807E3	44.773	7.552E3	50.569	6.807E3	41.034	8.282E3
	2	92.223		93.312		76.353		92.900	

The bold values are the largest values.

measure is proposed between the benchmark and the image after Gaussian filtering (to suppress random noises), to enhance true target and suppress background simultaneously. In the local contrast calculation, the DoG filter is adopted and improved with a non-negative constraint to get the weighting function to suppress clutters further. Experiments on 14 real sequences and a single frame dataset show the effectiveness of the proposed algorithm under different types of targets and backgrounds, and, compared to some baseline methods, the proposed algorithm can usually achieve better performance in SCRG, BSF and ROC curves. Besides, ablation experiments are conducted to verify the effectiveness of some important modules.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

JH: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. SM: Formal analysis, Writing – original draft, Writing – review & editing. WW: Methodology, Writing – original draft. NL: Data curation, Investigation, Writing – original draft. QZ: Project administration, Writing – original draft. ZL: Data curation, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported in part by the National Natural Science Foundation of

China under Grant 61802455, 62003381 and 62472464, in part by the Natural Science Foundation of Henan Province under Grant 252300420399, 242300421718 and in part by the Foundation of the Science and Technology Department of Henan Province under Grant 192102210089, 222102210077, 232102320066 and 252102210231.

Acknowledgments

Some of the testing IR sequences used in this paper are acquired from Hui et al. (2019) and Dai et al. (2021), we would like to acknowledge the authors for their kindly sharing. Also, We would like to express our sincere appreciation to the editor and reviewers who provided valuable comments to help improve this paper. Additionally, we are grateful to the researchers who provided comparative methods.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bai, X., and Zhou, F. (2010). Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognition* 43, 2145–2156. doi: 10.1016/j.patcog.2009.12.023
- Bi, Y., Chen, J., Sun, H., and Bai, X. (2020). Fast detection of distant, infrared targets in a single image using multiorder directional derivatives. *IEEE Trans. Aerospace Electronic Syst.* 56, 2422–2436. doi: 10.1109/TAES.2019.2946678
- Chen, C. L. P., Li, H., Wei, Y., Xia, T., and Tang, Y. Y. (2014). A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* 52, 574–581. doi: 10.1109/TGRS.2013.2242477
- Chen, Y., Song, B., Du, X., and Guizani, M. (2019). Infrared small target detection through multiple feature analysis based on visual saliency. *IEEE Access* 7, 38996–39004. doi: 10.1109/ACCESS.2019.2906076
- Chen, J., Zhu, Z., Hu, H., Qiu, L., Zheng, Z., and Dong, L. (2023a). Multi-scale local contrast fusion based on LoG in infrared small target detection. *Aerospace* 10, 449. doi: 10.3390/aerospace10050449
- Chen, J., Zhu, Z., Hu, H., Qiu, L., Zheng, Z., and Dong, L. (2023b). A novel adaptive group sparse representation model based on infrared image denoising for remote sensing application. *Appl. Sci.* 13, 5749. doi: 10.3390/app13095749
- Cui, Z., Yang, J., Jiang, S., and Li, J. (2016). An infrared small target detection algorithm based on high-speed local contrast method. *Infrared Phys. Technol.* 76, 474–481. doi: 10.1016/j.infrared.2016.03.023
- Dai, Y., and Wu, Y. (2017). Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 10, 3752–3767. doi: 10.1109/JSTARS.2017.2700023
- Dai, Y., Wu, Y., Song, Y., and Guo, J. (2017). Non-negative infrared patch-image model: Robust target background separation via partial sum minimization of singular values. *Infrared Phys. Technol.* 81, 182–194. doi: 10.1016/j.infrared.2017.01.009
- Dai, Y., Wu, Y., Zhou, F., and Barnard, K. (2021). Attentional local contrast networks for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* 59, 9813–9824. doi: 10.1109/TGRS.2020.3044958

- Dang, C., Li, Z., Hao, C., and Xiao, Q. (2023). Infrared small marine target detection based on spatiotemporal dynamics analysis. *Remote Sens.* 15, 1258. doi: 10.3390/rs15051258
- Deng, H., Sun, X., Liu, M., Ye, C., and Zhou, X. (2016). Infrared small-target detection using multiscale gray difference weighted image entropy. *IEEE Trans. Aerospace Electronic Syst.* 52, 60–72. doi: 10.1109/TAES.2015.140878
- Deng, H., Sun, X., Liu, M., Ye, C., and Zhou, X. (2017). Entropy-based window selection for detecting dim and small infrared targets. *Pattern Recognition* 61, 66–77. doi: 10.1016/j.patcog.2016.07.036
- Deng, L., Zhang, J., Xu, G., and Zhu, H. (2021). Infrared small target detection via adaptive m-estimator ring top-hat transformation. *Pattern Recognition* 112, 107729. doi: 10.1016/j.patcog.2020.107729
- Deshpande, S. D., Er, M. H., Venkateswarlu, R., and Chan, P. (1999). “Max-mean and max-median filters for detection of small targets,” in *1999 SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*. (Denver, CO, United States), 3809, 74–83. doi: 10.1117/12.364049
- Ding, H., and Zhao, H. (2015). Adaptive method for the detection of infrared small target. *Optical Eng.* 54, 113107. doi: 10.1117/1.OE.54.11.113107
- Du, P., and Hamdulla, A. (2020a). Infrared moving small-target detection using spatial-temporal local difference measure. *IEEE Geosci. Remote Sens. Lett.* 17, 1817–1821. doi: 10.1109/LGRS.2019.2954715
- Du, P., and Hamdulla, A. (2020b). Infrared small target detection using homogeneity-weighted local contrast measure. *IEEE Geosci. Remote Sens. Lett.* 17, 514–518. doi: 10.1109/LGRS.2019.2922347
- Fan, X., Wu, A., Chen, H., Huang, Q., and Xu, Z. (2022). Infrared dim and small target detection based on the improved tensor nuclear norm. *Appl. Sci.* 12, 5570. doi: 10.3390/app12115570
- Fang, H., Chen, M., Liu, X., and Yao, S. (2020). Infrared small target detection with total variation and reweighted l1 regularization. *Math. Problems Eng.* 2020, 1529704. doi: 10.1155/2020/1529704
- Gao, C., Meng, D., Yang, Y., Wang, Y., Zhou, X., and Hauptmann, A. G. (2013). Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* 22, 4996–5009. doi: 10.1109/TIP.2013.2281420
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Columbus, OH, USA: IEEE), 580–587. doi: 10.1109/CVPR.2014.81
- Gregoris, D. J., Yu, S. K., Tritchew, S., and Sevigny, L. (1994). “Wavelet transform-based filtering for the enhancement of dim targets in FLIR images,” in *Wavelet Applications*, vol. 2242. Ed. H. H. Szu (Orlando, FL, United States: International Society for Optics and Photonics (SPIE)), 573–583. doi: 10.1117/12.170058
- Guan, X., Peng, Z., Huang, S., and Chen, Y. (2020a). Gaussian scale-space enhanced local contrast measure for small infrared target detection. *IEEE Geosci. Remote Sens. Lett.* 17, 327–331. doi: 10.1109/LGRS.2019.2917825
- Guan, X., Zhang, L., Huang, S., and Peng, Z. (2020b). Infrared small target detection via non-convex tensor rank surrogate joint local contrast energy. *Remote Sens.* 12, 1520. doi: 10.3390/rs12091520
- Han, J., Liang, K., Zhou, B., Zhu, X., Zhao, J., and Zhao, L. (2018). Infrared small target detection utilizing the multiscale relative local contrast measure. *IEEE Geosci. Remote Sens. Lett.* 15, 612–616. doi: 10.1109/LGRS.2018.2790909
- Han, J., Liu, C., Liu, Y., Luo, Z., Zhang, X., and Niu, Q. (2021). Infrared small target detection utilizing the enhanced closest-mean background estimation. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 14, 645–662. doi: 10.1109/JSTARS.2020.3038442
- Han, J., Liu, S., Qin, G., Zhao, Q., Zhang, H., and Li, N. (2019). A local contrast method combined with adaptive background estimation for infrared small target detection. *IEEE Geosci. Remote Sens. Lett.* 16, 1442–1446. doi: 10.1109/LGRS.2019.2898893
- Han, J., Ma, Y., Huang, J., Mei, X., and Ma, J. (2016). An infrared small target detecting algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* 13, 452–456. doi: 10.1109/LGRS.2016.2519144
- Han, J., Ma, Y., Zhou, B., Fan, F., Liang, K., and Fang, Y. (2014). A robust infrared small target detection algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* 11, 2168–2172. doi: 10.1109/LGRS.2014.2323236
- Han, J., Moradi, S., Faramarzi, I., Liu, C., Zhang, H., and Zhao, Q. (2020). A local contrast method for infrared small-target detection utilizing a tri-layer window. *IEEE Geosci. Remote Sens. Lett.* 17, 1822–1826. doi: 10.1109/LGRS.2019.2954578
- Han, J., Xu, Q., Moradi, S., Fang, H., Yuan, X., Qi, Z., et al. (2022). A ratio-difference local feature contrast method for infrared small target detection. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2022.3157674
- Hao, C., Li, Z., Zhang, Y., Chen, W., and Zou, Y. (2024). Infrared small target detection based on adaptive size estimation by multidirectional gradient filter. *IEEE Trans. Geosci. Remote Sens.* 62, 1–15. doi: 10.1109/TGRS.2024.3502421
- Hao, X., Liu, X., Liu, Y., Cui, Y., and Lei, T. (2023). Infrared small-target detection based on background suppression proximal gradient and GPU acceleration. *Remote Sens.* 15, 5424. doi: 10.3390/rs15225424
- He, Y., Li, M., Zhang, J., and An, Q. (2015). Small infrared target detection based on low-rank and sparse representation. *Infrared Phys. Technol.* 68, 98–109. doi: 10.1016/j.infrared.2014.10.022
- Hui, B., Song, Z., Fan, H., Zhong, P., Hu, W., Zhang, X., et al. (2019). A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background. doi: 10.11922/sciedb.902
- Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 1254–1259. doi: 10.1109/34.730558
- Kong, X., Liu, L., Qian, Y., and Cui, M. (2016). Automatic detection of sea-sky horizon line and small targets in maritime infrared imagery. *Infrared Phys. Technol.* 76, 185–199. doi: 10.1016/j.infrared.2016.01.016
- Kou, R., Wang, C., Fu, Q., Yu, Y., and Zhang, D. (2022). Infrared small target detection based on the improved density peak global search and human visual local contrast mechanism. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 15, 6144–6157. doi: 10.1109/JSTARS.2022.3193884
- Kou, R., Wang, C., Peng, Z., Zhao, Z., Chen, Y., Han, J., et al. (2023). Small target segmentation networks: A survey. *Pattern Recognition* 143, 109788. doi: 10.1016/j.patcog.2023.109788
- Lin, Z., Ganesh, A., Wright, J., Wu, L., Chen, M., and Ma, Y. (2009). Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix.
- Liu, Y., Li, N., Cao, L., Zhang, Y., Ni, X., Han, X., et al. (2024). Research on infrared dim target detection based on improved YOLOv8. *Remote Sens.* 16, 2878. doi: 10.3390/rs16162878
- Liu, D., Li, Z., Liu, B., Chen, W., Liu, T., and Cao, L. (2017). Infrared small target detection in heavy sky scene clutter based on sparse representation. *Infrared Phys. Technol.* 85, 13–31. doi: 10.1016/j.infrared.2017.05.009
- Liu, L., Wei, Y., Wang, Y., Yao, H., and Chen, D. (2023). Using double-layer patch-based contrast for infrared small target detection. *Remote Sens.* 15, 3839. doi: 10.3390/rs15153839
- Lu, Z., Huang, Z., Song, Q., Bai, K., and Li, Z. (2022b). An enhanced image patch tensor decomposition for infrared small target detection. *Remote Sens.* 14, 6044. doi: 10.3390/rs14236044
- Lu, R., Yang, X., Li, W., Fan, J., Li, D., and Jing, X. (2022a). Robust infrared small target detection via multidirectional derivative-based weighted contrast measure. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2020.3026546
- Luo, Y., Li, X., Chen, S., Xia, C., and Zhao, L. (2022). IMNN-LWEC: A novel infrared small target detection based on spatial-temporal tensor model. *IEEE Trans. Geosci. Remote Sens.* 60, 1–22. doi: 10.1109/TGRS.2022.3230051
- Luo, Y., Li, X., Wang, J., and Chen, S. (2024). Clustering and tracking-guided infrared spatial-temporal small target detection. *IEEE Trans. Geosci. Remote Sens.* 62, 1–20. doi: 10.1109/TGRS.2024.3384440
- Ma, T., Yang, Z., Wang, J., Sun, S., Ren, X., and Ahmad, U. (2022). Infrared small target detection network with generate label and feature mapping. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2022.3140432
- Moradi, S., Moallem, P., and Sabahi, M. F. (2016). Scale-space point spread function based framework to boost infrared target detection algorithms. *Infrared Phys. Technol.* 77, 27–34. doi: 10.1016/j.infrared.2016.05.007
- Pang, D., Shan, T., Li, W., Ma, P., Tao, R., and Ma, Y. (2022a). Facet derivative-based multidirectional edge awareness and spatial-temporal tensor model for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15. doi: 10.1109/TGRS.2021.3098969
- Pang, D., Shan, T., Ma, P., Li, W., Liu, S., and Tao, R. (2022b). A novel spatiotemporal saliency method for low-altitude slow small infrared target detection. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2020.3048199
- Peng, L., Lu, Z., Lei, T., and Jiang, P. (2024). Dual-structure elements morphological filtering and local z-score normalization for infrared small target detection against heavy clouds. *Remote Sens.* 16, 2343. doi: 10.3390/rs16132343
- Qi, S., Ma, J., Li, H., Zhang, S., and Tian, J. (2014). Infrared small target enhancement via phase spectrum of quaternion fourier transform. *Infrared Phys. Technol.* 62, 50–58. doi: 10.1016/j.infrared.2013.10.008
- Qin, H., Han, J., Yan, X., Zeng, Q., Zhou, H., Li, J., et al. (2016). Infrared small moving target detection using sparse representation-based image decomposition. *Infrared Phys. Technol.* 76, 148–156. doi: 10.1016/j.infrared.2016.02.003
- Qin, Y., and Li, B. (2016). Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci. Remote Sens. Lett.* 13, 1890–1894. doi: 10.1109/LGRS.2016.2616416
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). “You only look once: Unified, real-time object detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Las Vegas, NV, USA: IEEE), 779–788. doi: 10.1109/CVPR.2016.91
- Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi: 10.1109/TPAMI.2016.2577031
- Shao, X., Fan, H., Lu, G., and Xu, J. (2012). An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system. *Infrared Phys. Technol.* 55, 403–408. doi: 10.1016/j.infrared.2012.06.001

- Tang, Y., Xiong, K., and Wang, C. (2023). Fast infrared small target detection based on global contrast measure using dilate operation. *IEEE Geosci. Remote Sens. Lett.* 20, 1–5. doi: 10.1109/LGRS.2023.3233958
- Wang, H., Hu, Y., Wang, Y., Cheng, L., Gong, C., Huang, S., et al. (2024). Infrared small target detection based on weighted improved double local contrast measure. *Remote Sens.* 16, 4030. doi: 10.3390/rs16214030
- Wang, X., Lv, G., and Xu, L. (2012). Infrared dim target detection based on visual attention. *Infrared Phys. Technol.* 55, 513–521. doi: 10.1016/j.infrared.2012.08.004
- Wang, X., Peng, Z., Kong, D., Zhang, P., and He, Y. (2017). Infrared dim target detection based on total variation regularization and principal component pursuit. *Image Vision Computing* 63, 1–9. doi: 10.1016/j.imavis.2017.04.002
- Wei, H., Ma, P., Pang, D., Li, W., Qian, J., and Guo, X. (2022). Weighted local ratio-difference contrast method for detecting an infrared small target against ground-sky background. *Remote Sens.* 14, 5636. doi: 10.3390/rs14225636
- Wei, Y., You, X., and Li, H. (2016). Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognition* 58, 216–226. doi: 10.1016/j.patcog.2016.04.002
- Wu, F., Zhang, T., Li, L., Huang, Y., and Peng, Z. (2024). “Rpcanet: Deep unfolding rpca based infrared small target detection,” in *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. (Waikoloa, HI, USA: IEEE). 4797–4806. doi: 10.1109/WACV57701.2024.00474
- Xiong, B., Huang, X., and Wang, M. (2021). Local gradient field feature contrast measure for infrared small target detection. *IEEE Geosci. Remote Sens. Lett.* 18, 553–557. doi: 10.1109/LGRS.2020.2976208
- Xu, K., Song, C., Xie, Y., Pan, L., Gan, X., and Huang, G. (2024). RMT-YOLOv9s: An infrared small target detection method based on UAV remote sensing images. *IEEE Geosci. Remote Sens. Lett.* 21, 1–5. doi: 10.1109/LGRS.2024.3484748
- Yang, P., Dong, L., and Xu, W. (2022). Detecting small infrared maritime targets overwhelmed in heavy waves by weighted multidirectional gradient measure. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2021.3080389
- Yang, W., and Shen, Z. (1998). Small target detection and preprocessing technology in infrared image sequences. *Infrared Laser Eng.* 27, 23–28.
- Yang, L., Yang, J., and Yang, K. (2004). Adaptive detection for infrared small target under sea-sky complex background. *Electron. Lett.* 40, 1083–1085. doi: 10.1049/el:20045204
- Yao, H., Liu, L., Wei, Y., Chen, D., and Tong, M. (2023). Infrared small-target detection using multidirectional local difference measure weighted by entropy. *Sustainability* 15, 1902. doi: 10.3390/su15031902
- Zhang, Z., Ding, C., Gao, Z., and Xie, C. (2023). ANLPT: Self-adaptive and non-local patch-tensor model for infrared small target detection. *Remote Sens.* 15, 1021. doi: 10.3390/rs15041021
- Zhang, X., Ding, Q., Luo, H., Hui, B., Chang, Z., and Zhang, J. (2019). Infrared small target detection based on an image-patch tensor model. *Infrared Phys. Technol.* 99, 55–63. doi: 10.1016/j.infrared.2019.03.009
- Zhang, Y., Li, Z., Siddique, A., Azeem, A., and Chen, W. (2024). A real-time infrared small target detection based on double dilate contrast measure. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 17, 16005–16019. doi: 10.1109/JSTARS.2024.3421646
- Zhang, X., Ren, K., Wan, M., Gu, G., and Chen, Q. (2017). Infrared small target tracking based on sample constrained particle filtering and sparse representation. *Infrared Phys. Technol.* 87, 72–82. doi: 10.1016/j.infrared.2017.10.003
- Zhang, P., Zhang, L., Wang, X., Shen, F., Pu, T., and Fei, C. (2021). Edge and corner awareness-based spatial-temporal tensor model for infrared small-target detection. *IEEE Trans. Geosci. Remote Sens.* 59, 10708–10724. doi: 10.1109/TGRS.2020.3037938
- Zhao, J., Tang, Z., Yang, J., and Liu, E. (2011). Infrared small target detection using sparse representation. *J. Syst. Eng. Electron.* 22, 897–904. doi: 10.3969/j.issn.1004-4132.2011.06.004
- Zhou, D., and Wang, X. (2023). Research on high robust infrared small target detection method in complex background. *IEEE Geosci. Remote Sens. Lett.* 20, 1–5. doi: 10.1109/LGRS.2023.3297523
- Zhu, H., Liu, S., Deng, L., Li, Y., and Xiao, F. (2020a). Infrared small target detection via low-rank tensor completion with top-hat regularization. *IEEE Trans. Geosci. Remote Sens.* 58, 1004–1016. doi: 10.1109/TGRS.2019.2942384
- Zhu, H., Ni, H., Liu, S., Xu, G., and Deng, L. (2020b). TNLRS: Target-aware non-local low-rank modeling with saliency filtering regularization for infrared small target detection. *IEEE Trans. Image Process.* 29, 9546–9558. doi: 10.1109/TIP.2020.3028457