Infrared Small Target Detection Based on the Improved Density Peak Global Search and Human Visual Local Contrast Mechanism

Renke Kou, Chunping Wang, Qiang Fu, Ying Yu, and Dongdong Zhang

***Abstract*—Effective detection of small targets plays a pivotal role in infrared (IR) search and track applications for modern military defense or attack. However, IR small targets are very difficult to detect because of their weak brightness, small size, and lack of shape, structure, texture, and other information elements. In order to simultaneously satisfy the robustness and timeliness of target detection, inspired by density peak clustering and the human visual system, an idea combining an improved density peak global search and local contrast calculation is proposed. First, the positions of candidate targets are determined in the preprocessed image using the improved density peak global search method (IDPGSM). Second, the saliency map is obtained using the double-weights enhanced local contrast method (DWELCM) for the candidate target neighborhood. Finally, adaptive threshold segmentation is used to detect IR small targets. Through comprehensive analysis of five evaluation indicators, the experimental results on seven real sequences and three hundred IR images of different scenes that the proposed method has better detection performance compared with six baseline methods. It can quickly and accurately determine the small target position in the case of severe background clutter and noise interference.**

***Index Terms*—Double-weights enhanced local contrast method (DWELCM), improved density peaks global search (IDPGSM), infrared (IR) small target, three-layer window.**

I. Introduction

An infrared (IR) detection system is a passive detection device that works in the IR spectrum of the atmospheric window band. It is an important part of advanced electronic warfare and integrated fire-control systems. In a complex and strong electromagnetic interference environment, an IR detection system has many advantages, such as strong anti-interference ability, good concealment, ability to work during the day and night, and high measurement accuracy. Therefore, it is widely used in weapon systems for IR early warning, reconnaissance, and precision guidance in airborne, shipborne, vehicle, and spaceborne applications [1]. In a complex battlefield environment, an IR detection system is often required to quickly locate and intercept targets at long distances. However, when the target is far from the detection system and the atmospheric attenuation is severe, the true targets occupy only a few pixels in the image field. Moreover, the shape, structure, and texture features of true targets are not obvious [2]. If the detection method is designed based on the characteristics of the target itself, it becomes more difficult than the surface target, and can easily cause false alarms and missed detections. Therefore, the essence of IR small-target detection is to make full use of the local and global features of the target to accurately separate the true targets from the complex background and random noise.

In recent decades, detection methods for IR small targets have mainly been divided into two categories: detection-before-tracking and tracking-before-detection. Tracking-before-detection methods must first obtain certain prior information for the targets, and the required calculation time is very long. Therefore, more attention has been paid to detection before tracking, which is a detection method for a single-frame IR image [3]. Among the current single-frame detection methods, the maximum median/mean filter [4], two-dimensional (2-D) least mean square filter (TDLMS) [5], mathematical morphology top-hot filter [6], and Butterworth filter [7] are widely used. However, the detection effect of these traditional methods is poor in the case of severe background clutter and low target signal-to-noise ratio (SNR). In the past ten years, inspired by the characteristics of the human visual system (HVS), the local contrast method has become a hot topic. In 2013, Chen et al. [8] proposed the local contrast method (LCM). Under this method, an eight-direction and double-window filter template was first constructed. The template was then traversed through each pixel of the image from left to right and top to bottom. Consequently, the computation required significant time, and noise and complex backgrounds were not suppressed. Based on the LCM, Han et al. [9] successively proposed a series of improved methods, such as improved LCM (ILCM), relative LCM (RLCM) [10], trilayer LCM (TLLCM) [2], and enhanced close-mean background estimation (ECMBE) [11]. The ILCM introduced the average value of the subblocks as a parameter, which can effectively suppress the influence of random noise points. However, when the target is small, it is easy to smooth it out. The RLCM is a difference-ratio-type local contrast method that can effectively suppress the complex background while enhancing the target. However, the calculation time required was too long. There were still many false positives

Manuscript received 3 May 2022; revised 22 June 2022; accepted 19 July 2022. Date of publication 27 July 2022; date of current version 5 August 2022. This work was supported by Equip Preresearch Project under Grant 301020205. *(Corresponding author: Qiang Fu.)*

The authors are with the Department of Electronic and Optical Engineering, Army Engineering University, Shijiazhuang 050003, China (e-mail: 365255860@qq.com; wcp@tom.com; 1245316750@qq.com; 569434537@qq.com; 1042911849@qq.com).

Digital Object Identifier 10.1109/JSTARS.2022.3193884

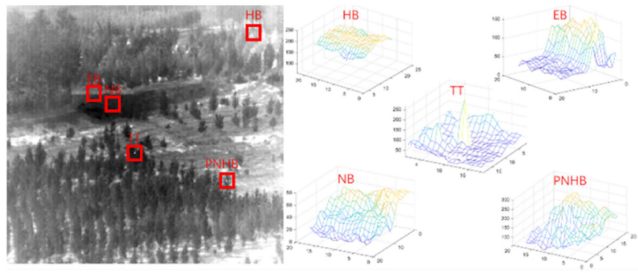
when the background was very complex. Based on the RLCM, the TLLCM performs Gaussian filtering on the center layer of the template, which effectively eliminates the influence of noise. The ECMBE method selects the pixel grey value closest to the core layer as the background estimation parameter. This method effectively overcomes situations where the target is easily overwhelmed owing to very bright backgrounds. However, this method can lead to several false alarms. Qin and Li [12] proposed a novel LCM (NLCM) that adopted a single-scale fixed template that was slightly larger than the target size. Although this method reduced the computation, the detection performance for complex backgrounds degraded because the estimation of the target size was not accurate. In addition, there are many improved local contrast methods that can be roughly divided into four categories. The first category is the difference type, such as the Laplacian of Gaussian (LoG) [13], difference of Gaussian (DoG) [14], improved Gabor [15], and multiscale patch-based contrast measure (MPCM) [16]. The second category is the ratio type, such as the weighted local difference measure [17]. The third category is the ratio-difference joint type, such as the homogeneity-weighted local contrast measure [18] and weighted strengthened local contrast measure [19]. The fourth category is a combination of traditional methods and local contrast methods, such as the multidirectional TDLMS [20].

The aforementioned series of methods inspired by the HVS face the following common problems.

1) These methods take the local features of the target as the research object and traverse the entire IR image through single-scale, multiscale (changing the window size), or image blocks. As a result, computational efficiency is restricted.
2) The detection results generate false alarms when the background clutter is severe and the target SNR is low.
3) These methods do not consider the global characteristics of the target in the IR image, which sometimes causes the missed detection of the true targets.

Therefore, in view of the abovementioned shortcomings, inspired by the density peak clustering and HVS, a new detection method is proposed in this article, which combines the improved density peak global search method (IDPGSM) and double-weights enhanced local contrast method (DWELCM).

The proposed method utilizes both global and local characteristics of the target. The innovations of this article are summarized as follows. For problems 1 and 3, according to the global characteristics of the target, a few candidate targets are determined by the IDPGSM. Then, DWELCM is performed on the candidate target neighborhood, which effectively reduces the amount of calculation, while ensuring its robustness. For problem 2, in the local contrast stage, the three-layer filtering window is improved. The core layer effectively overcomes the interference of random noise via matched filtering. Setting the middle layer effectively solves the impact of small changes in the target size. Double-weights can significantly enhance the target signal.



(a) (b)

Fig. 1. (a) Typical real IR image sample. (b) Three-dimensional mesh distributions of the different components.

The rest of this article is organized as follows. Section II presents an analysis of the characteristics of IR small targets. Section III describes the mathematical model of the proposed method in detail. Section IV demonstrates the robustness and timeliness of the proposed method from a mathematical point of view. Section V presents an analysis of the detection ability of the proposed method qualitatively and quantitatively from an experimental point of view. Finally, Section VI concludes this article.

II. Target Characteristic Analysis

A. Characteristics of IR Image Containing Small Target

As shown in Fig. 1, an IR single-frame image containing a small target consists of three parts: target, background, and noise. These three parts can be represented using an additive model [21]

$$f_I\ (x,y) = f_T\ (x,y) + f_B\,(x,y) + f_N\,(x,y) \qquad (1)$$

where (x, y) are the coordinates of each pixel in the image. f_I is the raw image. f_T, f_B, and f_N are the target, background, and noise components, respectively.

To explain the differences between the true target and various interferences intuitively, a typical sample of the real IR image is shown in Fig. 1(a). Five typical components are locally magnified, as shown in Fig. 1(b). The local area of the true target (TT) approximates a Gaussian distribution with a standard deviation of 29.67. The local area of the normal background (NB) is flat, and the gray value is generally low, with a standard deviation of 17.05. The local area of the high-brightness background (HB) is flat, and the overall gray value is high, with a standard deviation of 13.63. The local area of the edge background (EB) is stepped, with a standard deviation of 35.77. The local area of pixel-sized noise with high brightness (PNHB) has pixel-sized protrusions, with a standard deviation of 43.35.

This can be observed in Fig. 1. The standard deviations for PNHB, EB, and TT are relatively high. This is because the grayscale values of the pixels in these regions are relatively discrete. The standard deviations for NB and HB are relatively low. This is because these regions are relatively flat and have a small degree of dispersion. Therefore, the essence of IR small-target detection is to make full use of the difference between the target and nontarget features and accurately separate TT from the complex background (EB, HB, NB) and random PNHB.

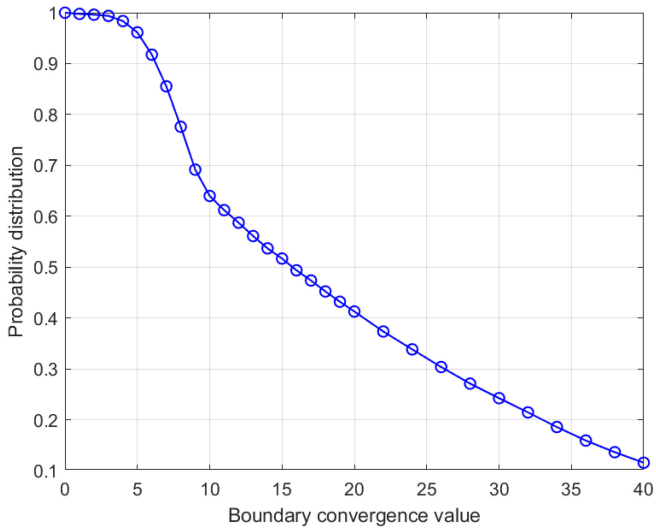


Fig. 2. Probability curve of target distribution.

B. Target Distribution Probability

To study the probability of the target distribution in the image field of view, we randomly select 10000 IR image samples from the literature [8]–[21] and dataset [22], [23]. Then, the TT within the sample is marked one by one. Finally, the raw image boundary is gradually reduced and the number of TT within the boundary are counted.

As shown in Fig. 2, 1) when the four boundaries of the raw image are shrunk inward by 1–3 pixel, the probability of the target appearing in the field of view of the image drops from 100% to 99.38%. This implies that the probability of a target at this boundary is very low. 2) When the boundary is reduced by 4–10 pixel, the probability of the target in the image field of view exhibits approximately exponential decay, which is sharply reduced from 98.36–63.99%. This implies that the target is relatively dense within the range of 4–10 pixel from the image boundary. 3) When the boundary is reduced by 11–40 pixel, the probability of the target in the image field of view decreases approximately linearly, which indicates that the distribution of the target in the middle area of the image field of view is random.

III. Proposed Method

The method proposed in this article includes four steps: 1) Preprocessed images are obtained using a DoG filter; 2) Candidate targets are determined using the IDPGSM; 3) The saliency map (SM) is obtained using DWELCM; 4) The TT position is determined using a threshold segmentation. The flow of the proposed method is shown in Fig. 3.

A. Identification of Candidate Targets by IDPGSM

1) Image Preprocessing: The DoG filter is a commonly used band pass filter that is widely used in image segmentation, edge detection, interest point detection, and other fields [14]. Therefore, according to the target feature analysis, this article uses the DoG filter to preprocess the raw IR image, which can enhance the target signal and suppress the background clutter.

The 1-D Gaussian function is defined as

$$G\ (r,\sigma\ ) = \frac{1}{\sqrt{2\pi}\sigma}\ e^{-\frac{r^2}{2\sigma^2}} \tag{2}$$

where σ is the standard deviation of the Gaussian function.

A 2-D Gaussian function, where the 1-D Gaussian function variable r is replaced by 2-D Gaussian function variable (i,j) can be defined as

$$G\ (i,j,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{i^2+j^2}{2\sigma^2}}. \tag{3}$$

Therefore, the DoG filter template is obtained by the difference between two Gaussian functions, which can be defined as

$$\begin{aligned}\text{DoG}\ (i,j,\sigma_1,\sigma_2)\ &=\ G\,(i,j,\sigma_1) - G\,(i,j,\sigma_2)\\ &= \frac{1}{2\pi}\ \left[\frac{1}{\sigma_1^2}e^{-\frac{i^2+j^2}{2\sigma_1^2}} - \frac{1}{\sigma_2^2}e^{-\frac{i^2+j^2}{2\sigma_2^2}}\right]\end{aligned} \tag{4}$$

where σ_1 and σ_2 are standard deviations of the Gaussian functions.

The raw image is convolved with the DoG filter template

$$I'=\ I * \text{DoG}\,(i,j,\sigma_1,\sigma_2) \tag{5}$$

where I is the raw image and I' is the preprocessed image.

2) Identification of Candidate Targets: Inspired by the density peak clustering method [24], the density (ρ) and δ-distance of each pixel in the IR image can be defined as

$$\rho\ = g_i \tag{6}$$

$$\delta_i = \min_{j:\rho_j>\rho_i}\ (d_{ij}) \tag{7}$$

$$d_{ij} = \sqrt{(x_i-x_j)^2+(y_i-y_j)^2} \tag{8}$$

where g_i is the gray value of pixel i, δ_i is the minimum distance (d_{ij}) between pixel i and any pixel with a higher density, (x_i,y_i) and (x_j,y_j) represent the coordinate positions of pixel *i* and *j*, respectively.

Because small targets have large ρ and δ values in the IR image, the density peak (p) can be defined as

$$p\ =\ \rho * \delta. \tag{9}$$

According to (9), the peak density value of each pixel in the preprocessed image can be obtained through traversal. Then, all density peaks are sorted in descending order, and the top n larger-density peaks are considered as candidate targets

$$\text{seeds}_n = \left\{(x,y)_{p_1},(x,y)_{p_2}\ldots(x,y)_{p_n}\right\} \tag{10}$$

where seeds_n is the set of the top *n* candidate target coordinate positions.

To further narrow the candidate targets range, the density peak search method is further improved. It can be observed in Fig. 2, that when the image boundary is indented by 3 pixels, the probability of the target still appearing in the image field of view is 99.38%. Therefore, setting the parameter l reasonably can not only ensure the detection probability but also reduce the

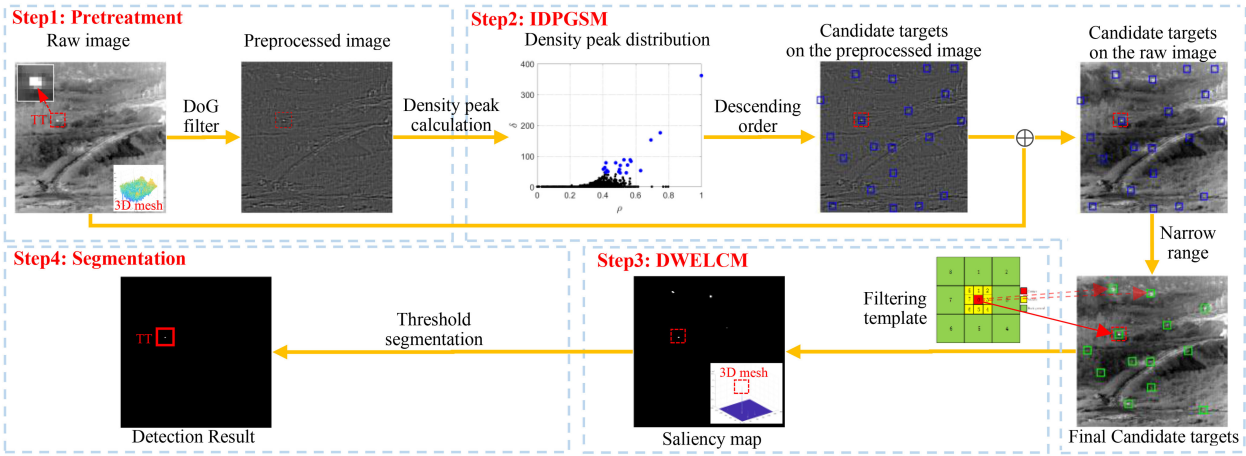


Fig. 3. Flowchart of the proposed method.

Algorithm 1: IDPGSM Computation.

Input: Single frame IR image (The total pixel size is H)
Output: The set of candidate target locations ($seeds_m$).
1: **for** $i = 1$ to H **do**
2: For the pixel i , obtain I' according to (5).
3: For the pixel i, obtain ρ and δ_i according to (6)–(8).
4: Compute p_i according to (9).
5: **end for**
6: Sort all the pixels by p_i in descending order according to (10). Record as $seeds_n$.
7: Adjust parameter l to reduce the number of candidates targets according to (11)–(12). Record as $seeds_m$.

number of candidate targets. The conditional function G can be defined as

$$G = \begin{cases} 0, \text{ if } x_i < l, \text{or } (M - x_i) < l \\ \text{ or } y_i < l, \text{or } (N - y_i) < l \\ 1, \qquad \text{otherwise} \end{cases} \tag{11}$$

$$\text{seeds}_m = \text{seeds}_n \cdot G \tag{12}$$

where M and N are the rows and columns of the IR image. l is the distance between the coordinates of the candidate targets and the image boundary. (x_i, y_i) are the coordinates of the candidate target i. seeds_m is the set of the top m candidate target coordinate positions. According to (11), m must be less than n.

The method for computing the IDPGSM is presented in Algorithm 1.

B. Target Detection by DWELCM

Inspired by the studies in [2] and [11], as shown in Fig. 4, a three-layer filtering template is improved, which includes the center, middle, and background layers. The center layer, which is 3×3 pixel, is used to capture the main energy of the target. The middle layer separates the target from the background and contains eight subblocks. The size of each subblock is 3×3 pixel.

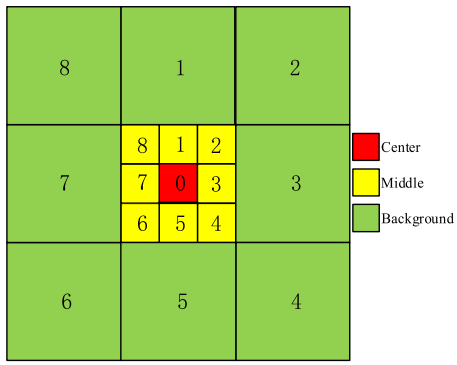


Fig. 4. Window template design.

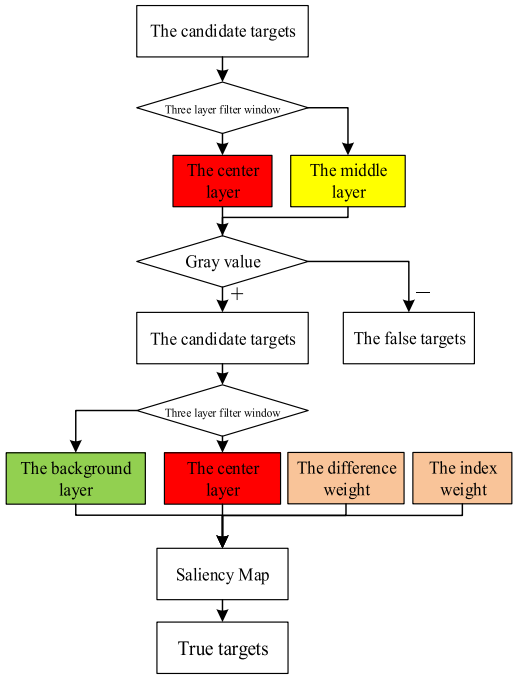


Fig. 5. Flow chart of DWELCM.

The background layer is used to estimate the background energy and has eight subblocks. The size of each subblock is 9×9 pixel.

The DWELCM process is shown in Fig. 5. First, the neighborhood of the candidate targets is traversed through an improved three-layer window template. Gaussian filtering is performed on the center layer of the candidate target area to further eliminate the influence of noise. Second, according to the characteristics of the target Gaussian distribution, the energy of the central layer must be greater than the average energy of the subblocks in the middle layer in any direction. The candidate targets are further screened in this manner. The average value of the top K2 largest pixel in each direction of the subblocks in the background layer is then calculated. Simultaneously, double-weights of the "exponential function" and "difference function" are introduced to further enhance the target signal. Finally, the location of the TT is determined using adaptive threshold segmentation. The mathematical modelling details are as follows:

1) Center Layer Gaussian Filter: Based on matched filter theory, the best filter to enhance a signal should have the same shape as the signal. The IR small-target image approximates an irregular 2-D Gaussian function distribution. Therefore, Gaussian filtering can be performed on the central layer to achieve the effect of enhancing the target and suppressing noise. The filter function is defined as follows:

$$I_{\text{center}} = \sum_{l=-1}^{1} \sum_{k=-1}^{1} GS\left(l,k\right) * I\left(x+l, y+k\right) \tag{13}$$

$$GS = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{14}$$

where (x, y) are the coordinates of the center layer. $I(\cdot)$ is the grey value of the raw image. I_{center} is the grey value after Gaussian filtering. GS is a Gaussian filter template.

2) Comparison of the Middle Layer and the Center Layer: To further filter out false targets from the n candidate targets, according to the Gaussian distribution characteristics of IR small targets, the gray value of the center layer must be greater than that of the middle layer in any direction. However, considering the existence of random noise, the average value of the top K1 largest pixels of the subblocks in the middle layer was taken as the comparison object

$$I_{\text{middle}}^{\text{mean}}\left(i\right) = \frac{1}{K1} \sum_{j=1}^{K1} I_{\text{middle}}^{j}\left(i\right) \tag{15}$$

$$L = \begin{cases} 1, \text{ if } I_{\text{core}} \geq \max\left[I_{\text{middle}}^{\text{mean}}\left(i\right)\right] \\ 0, \text{ otherwise} \end{cases} \tag{16}$$

where $I_{\text{middle}}^{j}(i)$ is the jth largest grey value in the ith subblock in the middle layer. $I_{\text{middle}}^{\text{mean}}(i)$ is the average value of the top K1 largest pixel of the ith subblock in the middle layer. L is the judgment criterion.

3) Background Layer Estimation: Similar to the calculation principle of the middle layer, to eliminate the influence of random noise, the average value of the top K2 largest pixel of the ith subblock in the background layer is introduced as the background estimation value, which is defined as

$$I_{\text{surround}}^{\text{mean}}\left(i\right) = \frac{1}{K2} \sum_{l=1}^{K2} I_{\text{surround}}^{l}\left(i\right) \tag{17}$$

where $I_{\text{surround}}^{l}(i)$ is the lth largest grey value in the ith subblock in the background layer. $I_{\text{surround}}^{\text{mean}}(i)$ is the average value of the top K2 largest pixel of the ith subblock in the background layer.

4) Double-Weights Enhanced Local Contrast Method: To enhance the target and suppress background clutter, according to (13)–(17), the enhanced local contrast can be defined as

$$\text{ELCM} = \min\left(\frac{I_{\text{center}}}{I_{\text{surround}}^{\text{mean}}\left(i\right)} \cdot I_{\text{center}} - I_{\text{center}}\right) \cdot L. \tag{18}$$

When the target SNR is low, the target enhancement effect is poor. To solve this problem, this article introduces two weights, namely the exponential weight W_1 and the difference weight W_2.

The target signal can be enhanced further based on the characteristics of the exponential function. W_1 is defined as

$$W_1 = \min\left(\exp\left(\frac{I_{\text{center}}}{I_{\text{surround}}^{\text{mean}}\left(i\right)} - 1\right)\right). \tag{19}$$

According to the characteristics of small targets, the smaller the mean difference of the gray between the center and background layers, the more is their similarity. To further enhance the target and suppress the background, the difference weight W_2 is introduced and defined as

$$I_{\text{center}}^{\text{mean}} = \frac{1}{9} \sum_{u=1}^{9} I_{\text{center}}^{u} \tag{20}$$

$$W_2 = \frac{1}{8} \sum_{i} \left|I_{\text{center}}^{\text{mean}} - I_{\text{surround}}^{\text{mean}}\left(i\right)\right| \tag{21}$$

where $I_{\text{center}}^{\text{mean}}$ is the mean of the center layer.

According to (18)–(21), the DWELCM can be defined as

$$\begin{aligned} \text{DWELCM} = &\min\left[\frac{I_{\text{center}}}{I_{\text{surround}}^{\text{mean}}\left(i\right)} \cdot \exp\left(\frac{I_{\text{center}}}{I_{\text{surround}}^{\text{mean}}\left(i\right)} - 1\right) - 1\right] \\ &\cdot I_{\text{center}} \cdot \frac{1}{8} \sum_{i} \left|I_{\text{center}}^{\text{mean}} - I_{\text{surround}}^{\text{mean}}\left(i\right)\right| \cdot L. \end{aligned} \tag{22}$$

The simplified expression is

$$\text{DWELCM} = \min\left[\frac{I_{\text{center}}}{I_{\text{surround}}^{\text{mean}}\left(i\right)} \cdot W_1 - 1\right] \cdot I_{\text{center}} \cdot W_2 \cdot L. \tag{23}$$

5) Threshold Segmentation: Through DWELCM calculations, among m candidate targets, the real target signal is amplified, and other candidate targets are suppressed. Finally, the TT position is determined using adaptive threshold segmentation. The threshold can be defined as

$$Th = \overline{\text{SM}} + k \cdot \sigma_{\text{SM}} \tag{24}$$

where $\overline{\text{SM}}$ is the mean value of the SM. σ_{SM} is the variance of SM. k is an empirical parameter.

In summary, the method for computing DWELCM is shown in Algorithm 2.

Algorithm 2: DWELCM Computation.

Input: Get the candidate target locations $seeds_m$ according to Algorithm 1

Output: The locations of TT.

1: **for** $seeds_m = 1$ to m **do**
2: The SM of candidate targets are obtained by DWELCM.
3: **end for**
4: The TT locations are determined by threshold segmentation according to (24).

IV. Mathematical Model Analysis

A. Robustness Analysis

1) If (x, y) is the TT, the gray value of TT is usually greater than the gray value of its neighbor pixel, and we can easily get

$$\begin{gathered}\min\left[\frac{I_{\text{center}}}{I^{\text{mean}}_{\text{surround}}(i)}\right](x,y) > 1\\ W_1(x,y) \gg 1\\ W_2(x,y) \gg 1\\ L = 1.\end{gathered}$$

Thus, $\text{WDELCM}_{\text{TT}}(x, y) \gg 1$.

2) If (x, y) is the NB or HB, we can easily get

$$\begin{gathered}\min\left[\frac{I_{\text{center}}}{I^{\text{mean}}_{\text{surround}}(i)}\right](x,y) \approx 0\\ W_1(x,y) \approx 0\\ W_2(x,y) \approx 0\\ L = 1 \text{ or } 0.\end{gathered}$$

Thus, $\text{WDELCM}_{\text{NB or HB}}(x, y) \approx 0$.

3) If (x, y) is the brighter or darker side of the EB, we can easily get

$$\begin{gathered}\min\left[\frac{I_{\text{center}}}{I^{\text{mean}}_{\text{surround}}(i)}\right](x,y) \approx 0\\ W_1(x,y) \geq 0\\ W_2(x,y) \geq 0\\ L = 1 \text{ or } 0.\end{gathered}$$

Thus, $\text{WDELCM}_{\text{EB}}(x, y) \approx 0$.

4) If (x, y) is the PNHB, we can easily get

$$\begin{gathered}\min\left[\frac{I_{\text{center}}}{I^{\text{mean}}_{\text{surround}}(i)}\right](x,y) \geq 1\\ W_1(x,y) \geq 0\\ W_2(x,y) \geq 0\\ L = 1.\end{gathered}$$

Thus, $\text{WDELCM}_{\text{PNHB}}(x, y) \geq 0$.

Although the noise is not completely suppressed, the target signal must be much larger than the noise

$$\text{WDELCM}_{\text{TT}} \gg \text{WDELCM}_{\text{PNHB}}.$$

B. Timeliness Analysis

The current local contrast calculation is mainly single-scale or multiscale (changing the window size to 3×3, 5×5, 7×7, 9×9) to traverse the entire image from left to right and top to bottom, or to introduce image blocks to reduce the amount of computation. However, this article first uses the IDPGSM to quickly determine m candidate targets. Then, only the m candidate target neighborhoods are calculated on a single scale, which significantly reduces the amount of computation. Taking an image with size 128×128 pixel as an example, the actual amount of computation involved in the local contrast calculation process is reduced by at least 10 times. The larger the raw image pixel, the more obvious the advantage of the proposed method. For a comparative analysis, please refer to the experimental section of Section V.

TABLE I
Details of the Test Sequences

Seq.	Image Size	Target Numbers	Target Size	Background Details
1	128×128	1	3×4	Sky-Cloud background; Heavy clutters; Heavy random noises.
2	172×172		4×5	Sky-Cloud background; Heavy clutters; Heavy random noises.
3	213×213		3×3	Jungle background; Bright halos
4	240×240		2×3	Sky-Cloud background; Heavy clutters; Heavy random noises.
5	256×256		3×3	Ground-Tree background; Heavy clutters
6	600×600	2	4×5 4×4	Sky-cloud-ground-green trees-building-crowd background;
7	600×600	3	3×3 3×3 3×4	Sky background; Bright halos; Heavy random noises.

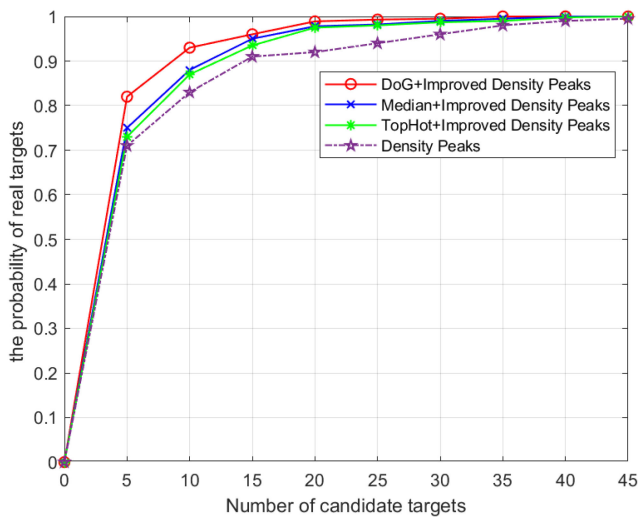


Fig. 6. Probability that candidate targets contain the TT.

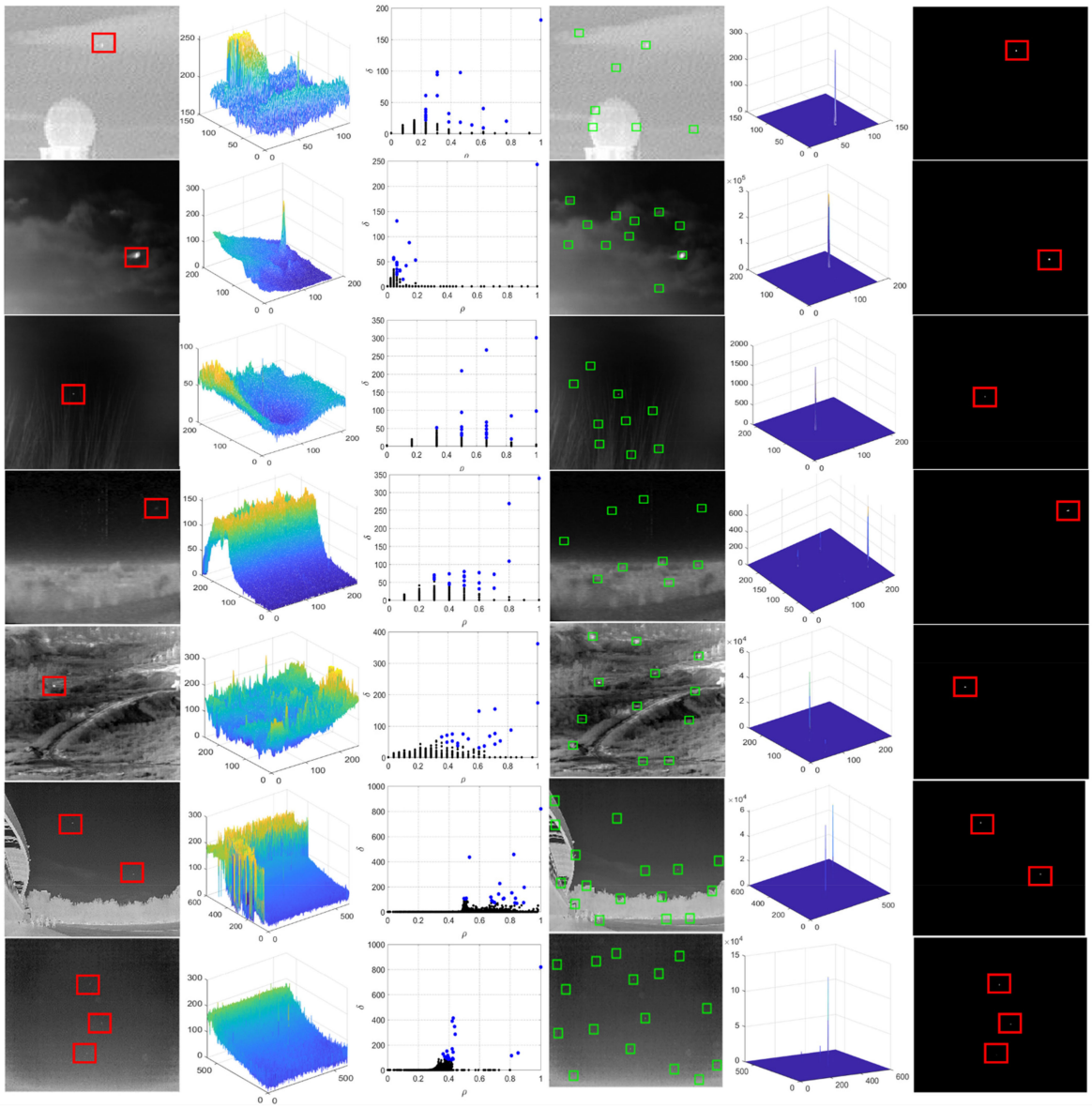


Fig. 7. Calculation process. (a) Raw IR images. (b) Three-dimensional mesh with Gaussian noise. (c) Density peak distribution. (d) Final candidate targets. (e) Saliency maps. (f) Detection results.

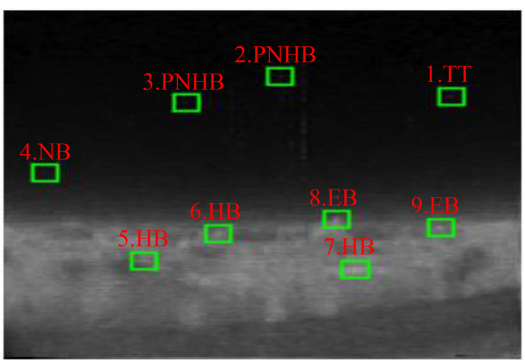


Fig. 8. Candidate targets are represented by green rectangles.

V. Experimental Results and Discussion

All experiments are conducted on a PC with a 2.3-GHz Intel i7-1180H CPU and 16.0-GB memory, and the code is implemented in MATLAB 2021a software. Here, σ_1 is 2; σ_2 is 8; n is 20; l is 3, K1 is 4; K2 is 9; k is 45.

A. Performance of the IDPGSM

In the stage of identifying candidate targets, in order to quantitatively compare the IDPGSM with density peak search algorithm proposed by Huang et al. [24], 1000 samples are randomly selected from the data set disclosed in [22] and [23] for testing. At the same time, in order to verify the quality of DoG filtering, median filtering and Top-hat filtering, a comparative

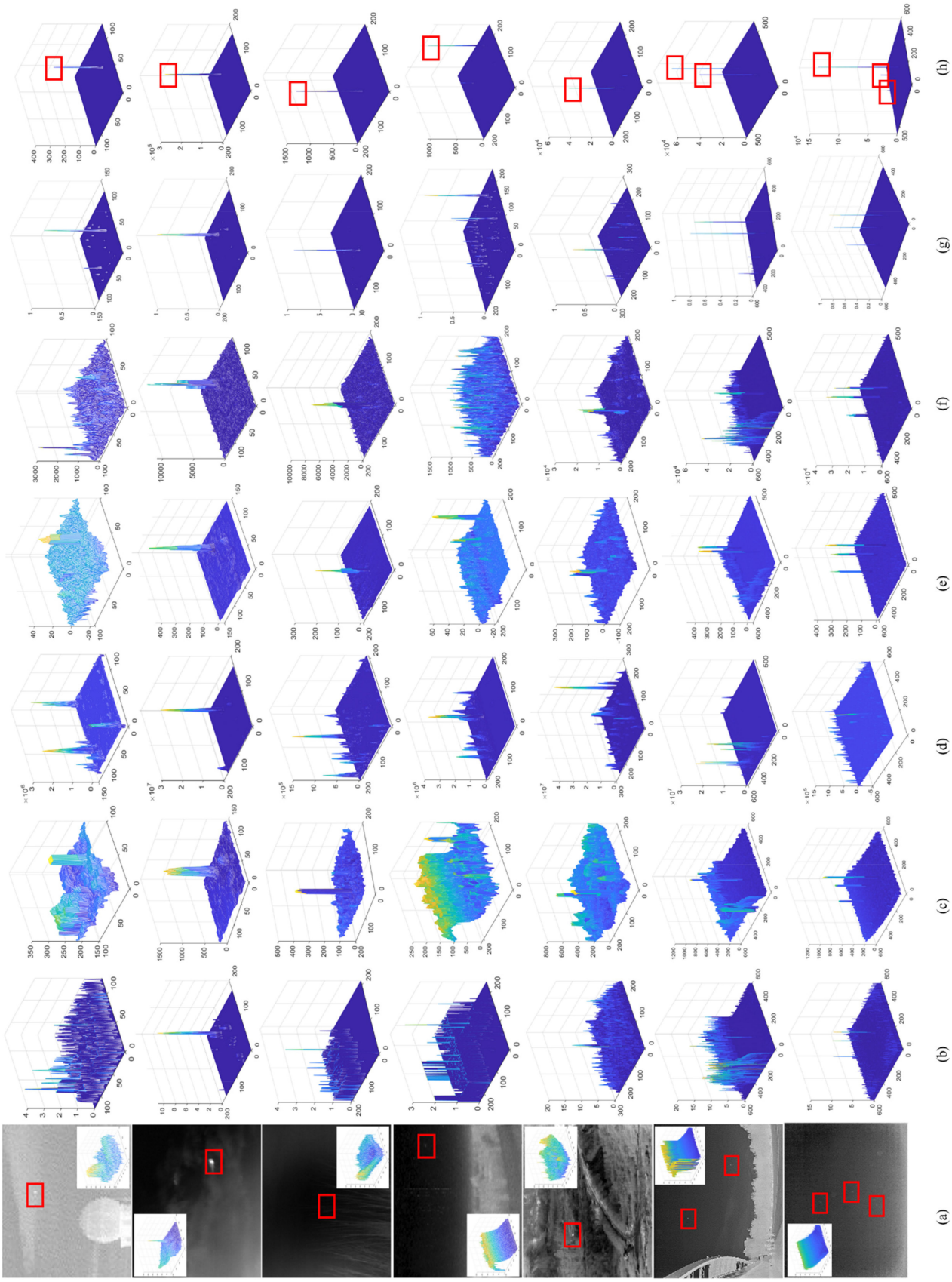


Fig. 9. Saliency maps of different methods. (a) Raw IR images. (b) DoG. (c) LCM. (d) MPCM. (e) RLCM. (f) NLCM. (g) TLLCM. (h) Proposed method.

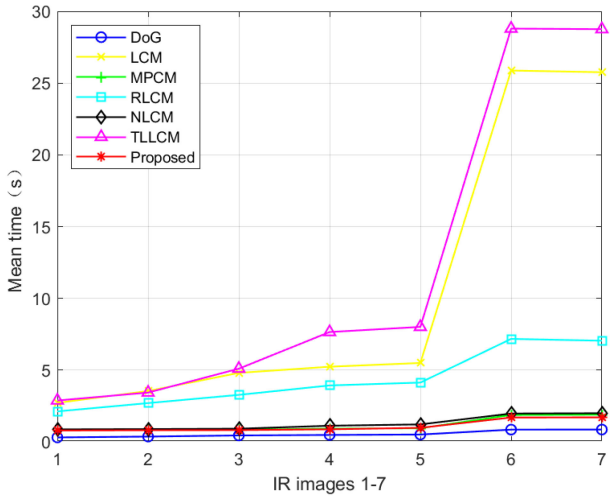


Fig. 10. Average calculation time.

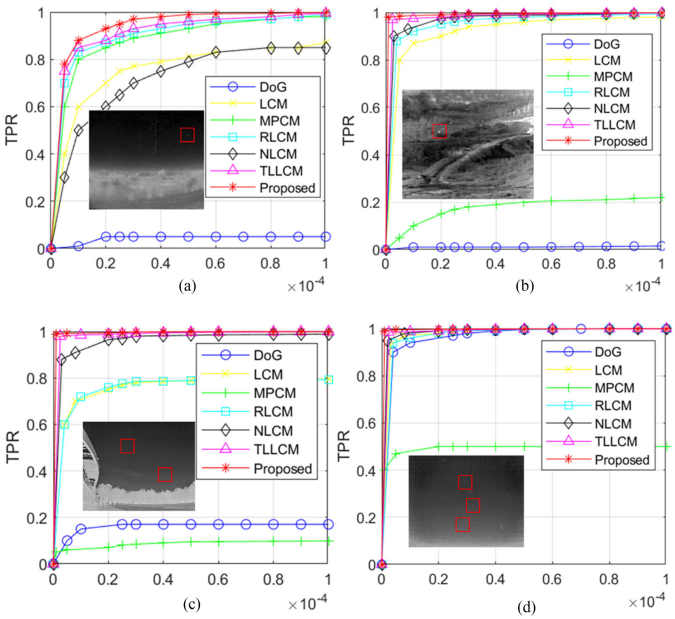


Fig. 11. ROC curve of IR images seq. 4–7.

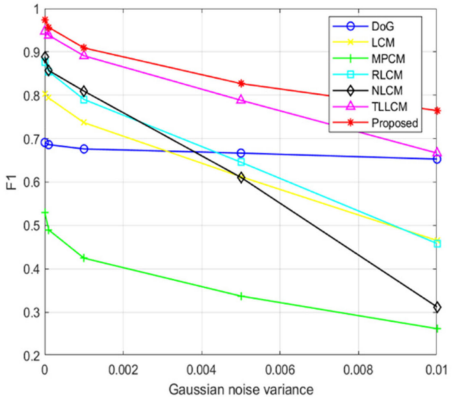


Fig. 12. Variation trend of F1 and noise.

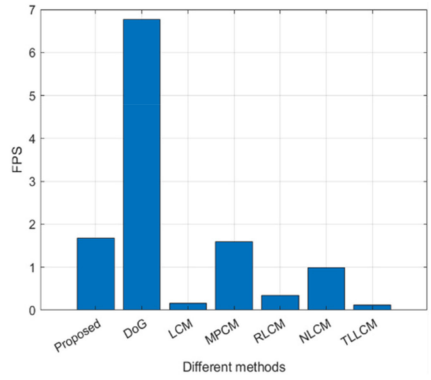


Fig. 13. FPS versus different methods.

experiment is carried out together. The experimental results are shown in Fig. 6.

It can be seen from Fig. 6 that the probability that the above-mentioned four algorithms can capture TT increases logarithmically with an increase in the number of candidate targets. But the density peak search method without preprocessing has the worst performance. Among the pretreatment methods, IDPGSM (the combination of DoG and improved density peak) has the best effect. When the candidate target n is set to 10, 15, 20, 25, and 30, the probabilities that the IDPGSM can capture TT are 93, 96, 98.9, 99.3, and 99.5%, respectively. Through a comparative analysis, it can be seen that under the same setting of candidate target n, the IDPGSM has a higher probability of capturing the TT.

B. Combined Performance of IDPGSM and DWELCM

To verify the robustness and timeliness of the proposed method under various complex IR backgrounds, five typical IR image sequences containing small targets are selected [10], [11], [24]. At the same time, in order to verify the multitarget detection ability of the proposed method, the IR images sequence of multi-uav small targets are collected by the long-wave (8–14 μm) IR detector. The details of these images are listed in Table I.

1) Robustness Analysis: The IR images in Table I are processed step-by-step according to the method proposed in this article. The test results are shown in Fig. 7. Fig. 7(a) shows the raw IR image. Fig. 7(b) shows the addition of Gaussian noise to the raw image (mean is 0, variance is 0.001). The purpose of adding random noise is to destroy the image quality and increase detection difficulty. Fig. 7(c) shows the distribution of the density peaks, in which 20 blue dots are the initially set candidate targets. Fig. 7(d) shows the final candidate targets after the IDPGSM calculation, which are represented by green rectangles. Fig. 7(e) shows the SM calculated using DWELCM. It can be observed that despite the interference of random noise and complex background, while the TT signal is enhanced, other background clutter and noise are suppressed. Fig. 7(f) shows the target location after threshold segmentation. It can be seen that through the method proposed in this article, small targets in the IR images of seven different scenes can all be detected.

TABLE II
GRAY VALUE OF CANDIDATE TARGETS

No.	1	2	3	4	5	6	7	8	9
Situation	TT	PNHB	PNHB	NB	HB	HB	EB	EB	EB
Gray value of the raw image	40	38	22	25	106	128	160	119	140
Gray value after filtering	**3971.47**	164.66	26.12	0	0	0	351.85	0	0

To further analyze the robustness of the mathematical model in this article, the nine candidate targets in the fourth row of the image in Fig. 7(d) are analyzed in detail, as shown in Fig. 8.

These candidate targets are categorized into five situations according to Fig. 1(a): TT, NB, HB, EB, and PNHB, as shown in Fig. 8. The gray values of nine candidate targets subsequent to the calculation using the proposed method are shown in Table II.

Through mathematical model analysis and the experimental results in Table II, it can be seen that if (x, y) is the TT, the pixel value is increased from 40 to 3971.47. If (x, y) is PNHB, the pixel value is much smaller than that of TT. If (x, y) is NB, HB, or EB, the pixel value is approximately equal to 0. The experimental results are in complete agreement with the analysis results of the mathematical model, further proving the effectiveness of the proposed method. In summary, the method proposed in this article is robust.

2) Comparisons With Other Methods: To verify the advantages of the method proposed in this article, six existing methods with high citation rates (DoG, LCM, MPCM, RLCM, NLCM, and TLLCM) are selected for comparative analysis. First, Gaussian noise (mean is 0, variance is 0.001) is added to the raw image. The SMs are then calculated sequentially using different methods, as shown in Fig. 9. As shown in Fig. 9(b), DoG filter has no directionality and is sensitive to background edges. Therefore, the final SM contains a large amount of background edge residue. It can be observed from Fig. 9(c) that LCM can enhance the target and noise simultaneously, and the background clutter suppression effect is not very good. As shown in Fig. 9(d), MPCM is a differential contrast method in which the background clutter is suppressed. However, interference is enhanced when the background is complex. As shown in Fig. 9(e), RLCM has a certain suppression effect on the noise and background, but when the background is complex, the suppression effect is general. As shown in Fig. 9(f), NLCM is similar to RLCM in single target detection, but it is better than RLCM in multitarget detection. As shown in Fig. 9(g), TLLCM has better suppression effect on the noise and background. As shown in Fig. 9(h), the method proposed in this article achieves good results in improving the target SNR and suppressing background and noise.

In Fig. 9, we can only qualitatively compare the advantages and disadvantages of the different methods from visual intuition. To conduct a quantitative analysis, according to [11], this article conducts a comparative analysis from five aspects: average calculation time, signal-clutter ratio gain (G_{SCR}), background suppression factor (B_{BSF}), ROC curve, and F1 Score.

a) Average calculation time: The test was repeated 5 times for each image and the average calculation time was taken. Fig. 10 shows the calculation times of the seven methods for the seven images in Table I (The image pixel size is shown in Table I). It can be seen that as follows:

1) LCM and TLLCM adopts a multiscale pixel-by-pixel calculation method, which leads to a long calculation time. With an increase in the number of image pixels, the calculation time increases exponentially.
2) RLCM introduces image block segmentation; therefore, the calculation time is less than that of the LCM.
3) NLCM adopts a single-scale window and introduces image block segmentation. Therefore, the calculation time of images 1–7 is smaller than that of RLCM.
4) MPCM adopts multiscale difference image blocks, and its computational efficiency is similar to that of NLCM.
5) DoG is the traditional filtering method with the highest computational efficiency, but its robustness is poor. Therefore, it is not suitable for small target detection.
6) The method proposed in this article has higher computational efficiency compared with other five methods (LCM, MPCM, RLCM, NLCM, and TLLCM).

b) G_{SCR} and B_{BSF}: To further compare the enhancement of small target signals and the suppression of background clutter using different methods, two evaluation indicators (G_{SCR} and B_{BSF}) [19] are introduced. The higher the G_{SCR} and B_{BSF}, the easier it is to detect the small target. The calculation formula is as follows:

$$S_{\text{SCR}} = \frac{|\mu_t - \mu_b|}{\sigma_b} \tag{25}$$

$$G_{\text{SCR}} = \frac{S_{\text{SCR\_out}}}{S_{\text{SCR\_in}}} \tag{26}$$

where μ_t and μ_b are the maximum grey value of the target neighborhood and the average grey value of the entire image, respectively. $S_{\text{SCR\_in}}$ is the signal-clutter ratio of the raw image. $S_{\text{SCR\_out}}$ is the signal-clutter ratio of the final SM

$$B_{\text{BSF}} = \frac{C_{\text{IN}}}{C_{\text{OUT}}} \tag{27}$$

where C_{IN} is the standard deviation of the raw images. C_{OUT} is the standard deviation of the final SM.

The calculation results of G_{SCR} and B_{BSF} of the 7 IR images in Table I are shown in Table III. It can be seen that compared with other five methods, TLLCM and proposed method in this article have relatively good effects on background suppression and target signal enhancement.

c) ROC curve analysis: In addition to the abovementioned three evaluation indicators, the true positive rate (TPR) and false positive rate (FPR) are also important indicators for evaluating

TABLE III
G_{SCR} , AND B_{BSF} OF DIFFERENT METHODS

Seq.	Target Numbers	Evaluation Indexs	Different Methods						
			DoG	LCM	MPCM	RLCM	NLCM	TLLCM	Proposed
1	1	B_{BSF}	5.421	23.478	0.784	13.457	2.475	3.548E3	**4.457E4**
		G_{SCR}	7.548	11.157	5.365	15.357	17.362	78.632	**138.418**
2	1	B_{BSF}	3.754	15.421	1.375	32.142	27.845	1.187E4	**3.784E4**
		G_{SCR}	15.653	21.625	11.235	20.736	37.003	113.485	**256.758**
3	1	B_{BSF}	0.984	47.487	0.475	17.473	0.876	5.865E4	**1.749E5**
		G_{SCR}	8.789	13.358	5.362	11.036	15.458	**237.041**	201.547
4	1	B_{BSF}	2.753	27.775	0.576	6.139	1.412	4.391E3	**2.478E4**
		G_{SCR}	6.288	7.847	7.397	15.785	10.189	102.532	**178.264**
5	1	B_{BSF}	0.574	8.475	0.257	2.491	0.874	3.004E3	**1.794E4**
		G_{SCR}	1.187	5.152	2.452	4.784	7.984	24.135	**78.457**
6		B_{BSF}	7.472	49.384	0.378	11.578	5.476	1.758E4	**2.748E5**
	1	G_{SCR}	3.475	7.452	1.762	22.352	34.632	115.462	**185.675**
	2		0.758	5.458	0.526	0.758	37.562	135.264	**201.316**
7		B_{BSF}	3.837	34.245	0.487	23.365	3.478	3.574E3	**5.789E4**
	1	G_{SCR}	7.475	8.457	2.758	18.362	32.854	278.706	**358.524**
	2		8.416	12.367	2.078	0.789	37.518	**235.684**	223.532
	3		5.895	11.125	3.625	0.968	29.047	**241.008**	178.395

the quality of the method. The formula used is as follows:

$$\mathrm{TPR} = \frac{\text{number of detected true targets}}{\text{total number of true targets}} \tag{28}$$

$$\mathrm{FPR} = \frac{\text{number of detected false targets}}{\text{total number of pixels in the whole image}}. \tag{29}$$

Four more complex IR image seq. 4–7 in Table I are selected as the experimental objects. The ROC curves are shown in Fig. 11. It can be clearly observed that the ROC curve of the proposed method in this article for the abovementioned four typical images is closer to the upper left corner and is relatively stable.

Through comprehensive analysis of average calculation time G_{SCR}, B_{BSF}, and ROC curves, it can be seen from Figs. 9–11, and Table III that as follows:

1) DoG filter is suitable for image segmentation and edge detection, so the performance of small target detection is the worst when the image has numerous background edge, as in Seq. 1, 3, 4, 5, and 6. However, when the background is clean, small target detection is fast and good, as shown in Figs. 10 and 11(d).
2) The B_{BSF} of MPCM is the lowest. Especially in the case of more complex background, some interference will be enhanced, as in Seq. 1, 5, and 6.
3) LCM is the most classical local contrast method, which is sensitive to noise. When there is a lot of random noise, the detection effect is not good, as shown in Fig. 11(a).
4) RLCM has better detection capability for single target images. However, as can be seen from Fig. 11(c), the ROC curve of RLCM is relatively low. This means that RLCM is not very good at multiobject image detection.
5) NLCM is similar to LCM and RLCM. However, NLCM calculation time is much faster than LCM or RLCM.
6) As can be seen from Table III and Fig. 11, the G_{SCR}, B_{BSF}, and ROC curves of TLLCM have good effects, which indicates that the method has good robustness. However, as can be seen from Fig. 10, TLLCM has a very large computation time, so it does not have real time.
7) The proposed method in this article has robustness and real-time performance compared with other six methods, as in Figs. 9–11, and Table III. Therefore, after comprehensive analysis, the proposed method in this article is more suitable for IR small target detection.

d) F1 Score and FPS: In order to further quantitatively evaluate the robustness of different methods, 300 IR images containing small targets in different scenes are selected (250 images are selected from the dataset [25], a single frame image is randomly selected from each of the 7 image sequences in Table I, and 43 images are actually taken).

Gaussian noise is added to each image in the test set (mean is 0, variance is 0, 0.0001, 0.001, 0.005, 0.01), and then seven methods are used to calculate true positive (TP, the number of detected true targets), false positive (FP, the number of detected false targets), and false negative (FN, the number of missed targets).

The precision and recall of target detection are calculated according to TP, FP, and FN, and the formula is

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{30}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}. \tag{31}$$

Where precison measures the consistency between prediction results and labels. Recall measures the ability to predict positive samples. Note: the statistical method of TP is that the distance between the centroid of the predicted target and true target is less than or equal to 4 pixels.

In order to comprehensively evaluate the robustness of the proposed method in this article, the F1 Score is introduced in

TABLE IV
TEST RESULTS OF TP, FP, FN, PRECISION, RECALL, F1 SCORE, AND FPS

Methods	Gaussian Noise (μ, σ^2)	TP	FP	FN	Precision	Recall	F1 Score	Times (FPS)
DoG	$(0, 0.0000)$	282	186	66	0.6025	0.8103	0.6911	**6.7751**
	$(0, 0.0001)$	282	192	66	0.5949	0.8103	0.6861	
	$(0, 0.0010)$	282	204	66	0.5802	0.8103	0.6762	
	$(0, 0.0050)$	282	216	66	0.5662	0.8103	0.6666	
	$(0, 0.0100)$	282	234	66	0.5465	0.8103	0.6527	
LCM	$(0, 0.0000)$	306	108	42	0.7391	0.8793	0.8031	0.1647
	$(0, 0.0001)$	300	108	48	0.7352	0.8620	0.7936	
	$(0, 0.0010)$	294	156	54	0.6533	0.8448	0.7368	
	$(0, 0.0050)$	270	264	78	0.5056	0.7758	0.6122	
	$(0, 0.0100)$	240	444	108	0.3508	0.6896	0.4651	
MPCM	$(0, 0.0000)$	288	450	60	0.3902	0.8275	0.5303	1.5918
	$(0, 0.0001)$	282	522	66	0.3507	0.8103	0.4895	
	$(0, 0.0010)$	270	654	78	0.2922	0.7758	0.4245	
	$(0, 0.0050)$	270	984	78	0.2153	0.7758	0.3370	
	$(0, 0.0100)$	264	1404	84	0.1582	0.7586	0.2619	
RLCM	$(0, 0.0000)$	318	60	30	0.8412	0.9137	0.8760	0.3479
	$(0, 0.0001)$	318	78	30	0.8030	0.9137	0.8548	
	$(0, 0.0010)$	294	102	54	0.7424	0.8448	0.7903	
	$(0, 0.0050)$	246	168	102	0.5942	0.7068	0.6456	
	$(0, 0.0100)$	198	318	150	0.3837	0.5689	0.4583	
NLCM	$(0, 0.0000)$	312	42	36	0.8813	0.8965	0.8888	0.9884
	$(0, 0.0001)$	306	60	42	0.8360	0.8793	0.8571	
	$(0, 0.0010)$	306	102	42	0.7500	0.8793	0.8095	
	$(0, 0.0050)$	240	198	108	0.5479	0.6896	0.6106	
	$(0, 0.0100)$	132	366	216	0.2650	0.3793	0.3120	
TLLCM	$(0, 0.0000)$	324	12	24	0.9642	0.9310	0.9473	0.3479
	$(0, 0.0001)$	318	12	30	0.96364	0.9137	0.9380	
	$(0, 0.0010)$	294	18	54	0.9423	0.8448	0.8909	
	$(0, 0.0050)$	246	30	102	0.8913	0.7068	0.7884	
	$(0, 0.0100)$	198	48	150	0.8048	0.5689	0.6666	
Proposed	$(0, 0.0000)$	330	0	18	1.0000	0.9482	**0.9734**	**1.6750**
	$(0, 0.0001)$	330	12	18	0.9649	0.9482	**0.9565**	
	$(0, 0.0010)$	300	12	48	0.9615	0.8620	**0.9090**	
	$(0, 0.0050)$	258	18	90	0.9347	0.7413	**0.8269**	
	$(0, 0.0100)$	234	30	114	0.8863	0.6724	**0.7647**	

this section, and the formula is

$$\mathrm{F1} = \frac{2 * \mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{32}$$

where F1 Score combines the results of precision and recall.

The F1 Score will be high only when both the Precision and Recall values are relatively large. And the higher the F1 Score, the more effective the model is. At the same time, the frames per second (FPS) of different methods are counted on the test set.

All the calculation results (TP, FP, FN, Precision, Recall, F1 Score, and FPS) are shown in Table IV.

It can be seen from Table IV that as the variance of Gaussian noise increases, the precision and recall of the detection methods both decrease. However, in the case of adding the same noise, the method proposed in this article has the highest F1 Score compared to the other six methods. This shows that our method is more robust. In order to more intuitively show the change trend of F1 Score and FPS of different methods, the F1 Score curve and FPS histogram are drawn, as shown in Figs. 12 and 13.

As can be seen from Fig. 12, the F1 curve of DoG is the flattest and insensitive to noise, but the F1 score is not high. The MPCM has the lowest F1 Score, so its robustness performance is

the worst. The F1 Score of NLCM, LCM, and RLCM decrease rapidly with the increase of noise, so the antinoise performance is not very good. With the noise increases, the F1 score of TLLCM decreases gently and is generally higher, so the robustness is better. The proposed method in this article achieves an F1 score of 97.34% for IR images of different scenes without noise interference. And with the increase of noise, the decrease of F1 score is more gradual and higher than the other six methods.

As can be seen from Fig. 13, the FPS of our method ranks second, only lower than DoG. But the F1 Score of our method is much higher than DoG.

In summary, our method balances robustness and timeliness. Therefore, it is more suitable for IR small target detection.

VI. Conclusion

In this article, an IR small-target detection method combining the density peak global search and HVS contrast mechanism is proposed for the first time. It consists of two core steps: the first is determining the candidate target location using IDPGSM, while the second is determining the location of the TT using DWELCM. Through quantitative and qualitative analyses of the mathematical model and experiment, the following results were obtained: 1) The IDPGSM improves the detection probability, and also reduces the candidate targets range. 2) Compared with the current local contrast methods with a high citation rate, the combination of IDPGSM and DWELCM greatly improves the computational efficiency as well as greatly enhances the TT signal. Moreover, it has a better suppression effect on background clutter and noise.

Acknowledgment

The authors would like to thank Prof. Z. Peng of University of Electronic Science and Technology and also like to thank Lecturer J. Han of Zhoukou Normal University for providing some source codes.

References

[1] R. Kou, H. Wang, Z. Zhao, and F. Wang, "Optimum selection of detection point and threshold noise ratio of airborne infrared search and track systems," *Appl. Opt.*, vol. 56, no. 18, pp. 5268–5273, Jun. 2017.

[2] J. Han, S. Moradi, I. Faramarzi, C. Liu, H. Zhang, and Q. Zhao, "A local contrast method for infrared small-target detection utilizing a tri-layer window," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1822–1826, Oct. 2020.

[3] S. S. Rawat, S. K. Verma, and Y. Kumar, "Review on recent development in infrared small target detection algorithms," *Procedia Comput. Sci.*, vol. 167, pp. 2496–2505, Apr. 2020.

[4] J. F. Rivest and R. Fortin, "Detection of dim targets in digital infrared imagery by morphological image processing," *Opt. Eng.*, vol. 35, no. 7, pp. 1886–1894, 1996.

[5] T. Soni, J. R. Zeidler, and W. H. Ku, "Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 327–340, Jul. 1993.

[6] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.,* vol. 43, no. 6, pp. 2145–2156, Jan. 2010.

[7] L. Yang, J. Yang, and K. Yang, "Adaptive detection for infrared small target under sea-sky complex background," *Electron. Lett.*, vol. 40, no. 17, pp. 1083–1085, Aug. 2004.

[8] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.

[9] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.

[10] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.

[11] 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 J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 645–662, 2021, doi: 10.1109/JSTARS.2020.3038442.

[12] Y. Qin and B. Li, "Effective infrared small target detection utilizing a novel local contrast method," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1890–1894, Dec. 2016.

[13] S. Kim, Y. Yang, J. Lee, and Y. Park, "Small target detection utilizing robust methods of the human visual system for IRST," *J. Infrared Millimeter Terahertz Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.

[14] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infrared Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.

[15] J. Han, Y. Ma, J. Huang, X. Mei, and J. Ma, "An infrared small target detecting algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 452–456, Mar. 2016.

[16] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for infrared small target detection," *Pattern Recognit.,* vol. 58, pp. 216–226, Oct. 2016.

[17] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Infrared small target detection based on weighted local difference measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4204–4214, Jul. 2016.

[18] P. Du and A. Hamdulla, "Infrared small target detection using homogeneity-weighted local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 514–518, Mar. 2020.

[19] J. Han et al., "Infrared small target detection based on the weighted strengthened local contrast measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 18, no. 9, pp. 1670–1674, Sep. 2021.

[20] J. Han, S. Liu, G. Qin, Q. Zhao, H. Zhang, and N. Li, "A local contrast method combined with adaptive background estimation for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1442–1446, Sep. 2019.

[21] 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 Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.

[22] B. Hui et al., "A dataset for infrared detection and tracking of dim-small aircraft targets under ground /air background," *Chin. Sci. Data*, vol. 5, no. 3, pp. 291–302, Aug. 2020.

[23] L. Huang, S. Dai, T. Huang, X. Huang, and H. Wang, "Infrared small target segmentation with multiscale feature representation," *Infrared Phys. Technol.*, vol. 116, May 2021, Art. no. 103755.

[24] S. Huang, Z. Peng, Z. Wang, X. Wang, and M. Li, "Infrared small target detection by density peaks searching and maximum-gray region growing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1919–1923, Dec. 2019.

[25] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Attentional local contrast networks for infrareds small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9813–9824, Nov. 2021.

Renke Kou received the B.E. degree in mechanical and electronic engineering from Changchun University of Technology, Changchun, China, and the M.E. degree in weapon science and technology from Air Force Engineering University, Xi 'an, China, in 2015 and 2017, respectively. He is currently working toward the Ph.D. degree with the department of electronic and optical engineering, Army Engineering University, Shijiazhuang, China.

His research interests include image process and pattern recognition.

Chunping Wang received the Ph.D degree in electronic and optical engineering from Shijiazhuang Mechanical Engineering College, Shijiazhuang, China, in 1999.

He is currently a Professor with Army Engineering University, Shijiazhaung, China. He has (co-authored 200 papers. He obtained four national defense invention patents. His research interest includes control theory and application, information processing.

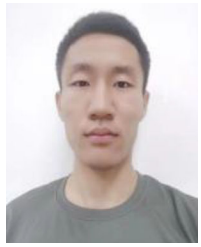
Prof. Wang was the recipient of the military science and technology progress award. At the same time, he is the editorial board member of *the journal of Fire Control and Command Control, and Aerodynamic Missile.*

Qiang Fu received the B.S. degree in electrical engineering from Shijiazhuang Mechanical Engineering College, Shijiazhuang, China, in 1999, and the Ph.D degree in computer science and technology from Tsinghua University, Beijing, China, in 2017.

He is a Lecturer with the Army Engineering University, Nanjing, China. He has authored more than 50 technical papers. His teaching and research interests include automatic control and image engineering.

Ying Yu received the B.E. degree in computer science and technology from Hebei Normal University, Shijiazhuang, China, in 2011, and the M.S. degree in computer software and theory from Yunnan Normal University, Kunming, China, in 2014. She is currently working toward the Ph.D. degree with the department of electronic and optical engineering, Army Engineering University, Shijiazhuang, China.

From 2014 to 2022, she was a Faculty with the School of Information and Intelligence Engineering, University of Sanya, Sanya, China, where she has been an Associate Professor since 2021. Her research interests include computer vision and pattern recognition.



Dongdong Zhang received the B.E. degree in radar engineering in 2020 from Army Engineering University, Shijiazhuang, China, where he is currently working toward the master's degree in ship target detection in SAR imagery.

His research interests include image processing and target recognition.