

Small Infrared Target Detection Based on Fast Adaptive Masking and Scaling With Iterative Segmentation

Yaohong Chen, Gaopeng Zhang, Yingjun Ma, Jin U. Kang, *Member, IEEE*, and Chiman Kwan 

Abstract—Fast and robust small infrared (IR) target detection is a challenging task and critical to the performance of IR searching and tracking (IRST) systems. However, the current algorithms generally have difficulty in striking a good balance between speed and performance. In this letter, we propose a new approach to small IR target detection that can significantly accelerate the detection process by first performing a fast adaptive masking and scaling algorithm. We then propose to enhance the target characteristics and suppress the background clutter using both contrast and gradient information. Finally, we propose to accurately extract the targets via iterative segmentation. The experimental results demonstrated that our proposed method yields the best and the most robust performance, with a speed of at least two times faster than the state-of-the-art methods.

Index Terms—Adaptive masking and scaling, iterative segmentation, real-time target detection, small infrared (IR) target detection.

I. INTRODUCTION

THE infrared searching and tracking (IRST) system has been widely used in defense and civilian applications [1]. The operation distance of the IRST system is highly related to the detection performance of small targets, which are typically in complex backgrounds and with low signal-to-clutter ratio (SCR) and small size [2]. Consequently, small infrared (IR) target detection is a challenging task that has become increasingly significant in the past few decades.

The single-frame detection methods are widely used due to their flexibility and low computational cost [2]. Filtering-based methods such as Top-hat filter [3] and Max-mean filter [4] cannot achieve promising results in complex backgrounds.

Manuscript received June 5, 2020; revised August 11, 2020 and November 2, 2020; accepted December 21, 2020. This work was supported by the National Natural Science Foundation of China under Grant 51905529. (Corresponding authors: Jin U. Kang; Chiman Kwan.)

Yaohong Chen is with the Xian Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China, also with the University of Chinese Academy of Sciences, Beijing 100049, China, and also with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: ychen405@jhu.edu).

Gaopeng Zhang and Yingjun Ma are with the Xian Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China (e-mail: mayingjun@opt.ac.cn; zhanggaopeng@opt.ac.cn).

Jin U. Kang is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: jkang@jhu.edu).

Chiman Kwan is with the Signal Processing, Inc., Rockville, MD 20850 USA (e-mail: chiman.kwan@signalpro.net).

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/LGRS.2020.3047524>.

Digital Object Identifier 10.1109/LGRS.2020.3047524

Blob detection methods, including the Laplacian of Gaussian (LoG) filter [5] and the difference of Gaussian (DoG) [6] filter, are susceptible to flickering noise and spot-like backgrounds. Chen *et al.* [7] proposed to detect small IR targets using local contrast measure (LCM), which has difficulty in distinguishing targets and strong clutters. To improve the performance of LCM, Wei *et al.* [8] introduced a multiscale patch-based contrast measure (MPCM). Gao *et al.* [9] developed an IR patch-image (IPI) model to convert small target detection into an optimization problem. Zhang *et al.* [10] improved the performance of the IPI via nonconvex rank approximation minimization (NRAM). Recently, Zhang *et al.* [11] proposed to detect small IR targets based on the local intensity and gradient (LIG) properties. Qin *et al.* [12] and Zhang *et al.* [13] developed a fast algorithm based on facet kernel and random walker (FKRW) and partial sum of the tensor nuclear norm (PSTNN), but these methods have difficulty in achieving real-time processing in complex scenarios.

According to [5], [7], and [11], small IR targets have several important characteristics. First, the small targets exhibit pop-out behavior, meaning that the target textures and SCR are quite distinct from their surrounding pixels. Second, the target pixels are different from the background pixels in terms of local contrast and gradient values. The combination of pop-out behavior and local SCR plays a key role in searching for target candidates efficiently, while the local contrast and gradient properties are effective for accurately distinguishing real targets from strong background clutters.

In this letter, we propose a novel IR small target detection algorithm known as fast adaptive masking and scaling with iterative segmentation (FAMSIS). Our method consists of masking, adaptive scale factor estimation, contrast and gradient map generation, and target detection using iterative segmentation. The main contributions of this letter are summarized as follows:

- 1) We propose a novel and real-time framework for small target detection by comprehensively and simultaneously incorporating target size, local contrast, and gradient.
- 2) The experimental results demonstrate that the proposed method runs at least two times faster than the state-of-the-art methods [12] without degrading the detection performance.

II. PROPOSED METHOD

Fig. 1 shows the architecture of our proposed method. In Phase 1, we first apply a DoG filter to the input image and calculate a binarized mask, and adaptively estimate a scale

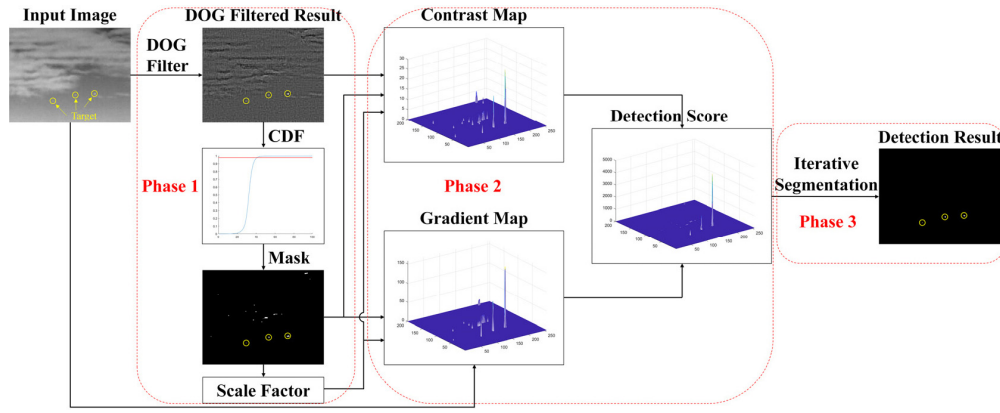


Fig. 1. Framework diagram of the proposed method.

factor based on the mask. The masks and scale factors enable a fast and accurate overall target detection system. In Phase 2, we combine the local contrast and gradient properties for high-performance target detection and background suppression. Finally, we accurately extract the targets by performing iterative segmentation.

A. Phase 1: Masking and Scale Factor Estimation

Generally, an IR image with small targets consists of background, target, and noise. Since the spatial frequencies of the targets are typically higher than the background and lower than the noise [6], it is clear that the DoG filter, which is widely used as a bandpass filter [14], can effectively extract salient target candidates and suppress background clutter and noise. The DoG filter can be defined by [6]

$$\text{DoG}(i, j, \sigma_1, \sigma_2) = \frac{1}{2\pi\sigma_1^2} e^{-(i^2+j^2)/2\sigma_1^2} - \frac{1}{2\pi\sigma_2^2} e^{-(i^2+j^2)/2\sigma_2^2} \quad (1)$$

where σ_1 and σ_2 are the customer-selected parameters that represent the standard deviations of Gaussian functions. In the case of $\sigma_2/\sigma_1 = 1.6$, the DoG filter can be seen as an approximation of LoG filter [14], which is one of the most common differential blob detectors. Equation (1) is mathematically equivalent to

$$\text{DoG}(i, j, \sigma_1, \sigma_2) = \text{DoG}(i, j, \sigma_1, \sigma_1) + \text{DoG}(i, j, \sigma_1, \sigma_2) \quad (2)$$

which can be regarded as a concatenation of two LoG filters if $\sigma_2 = 1.6 \times \sigma_1 = 1.6 \times 1.6 \times \sigma_1$. Given that the sidelobe of a Laplacian filter crosses zero with a target diameter of $2\sqrt{2}\sigma$ [5], and the target sizes in our data sets range from 3×3 to 7×7 , the smaller standard deviation σ_1 of the DoG filter is defined by the smallest target size. Consequently, $\sigma_1 = 3/2\sqrt{2} = 1.06$, while σ_2 is set to 2.71 according to $\sigma_2 = 1.6 \times 1.6 \times \sigma_1$, and a minor adjustment of σ_2 will not significantly affect the results.

Let I_D denote the processing result of the DoG filter. It is clear that the target candidates, including real targets and background clutter, are all enhanced with a higher intensity than other components. We calculate the cumulative distribution function (CDF) of I_D and extract the pixels with the highest intensity to generate a binarized mask I_M . We regard the percentage of candidate pixels as p , which is an adjustable parameter and a value of $p = 0.2\%$ works well in all our experiments. Similar to past articles, it should be emphasized

Algorithm 1 Adaptive Scale Factor Estimation

Input: Binarized mask I_M .

Output: Scale factor s for each connected region in I_M .

```

1: Detect connected regions and calculate the number of
   regions ( $N_C$ ).
2: for  $c = 1: N_C$  do
3:   Generate minimum binarized matrix  $L$  containing
   this region.
4:   for  $m = 2: m_{\max}$  do
5:      $P_m = L \cdot O_m$ .
6:     if  $P_m > m \times (m-1)$  do
7:        $s_c = m$ .
8:     end if
9:   end for
10: end for

```

that we only focus on detecting bright targets, and a low CDF cutoff threshold is required for the detection of dark targets. It should be emphasized that all the subsequent steps process only those pixels represented by the mask.

In [10] and [11], the scale factors are related to the target sizes and have been applied to improve the target detection performance. Motivated by the above finding, we propose to calculate adaptive scale factors for target candidates in the masks obtained by the DoG filter. First, we detect all the connected nonzero regions and isolated nonzero pixels in the masks. For these isolated nonzero pixels, we set the scale factor s to 3 in the subsequent processing. Second, we apply all-ones matrix O_m with a size of $m \times m$ to each connected region, with m ranging from 2 to the maximum morphological size of the connected region. The estimation for the adaptive scale factor s of each region is illustrated in Algorithm 1. It should be noted that for the connected regions with a scale factor less than 3, we define the scale factor to be 3, and for the scale factor with an even value, we adjust it to an odd integer by adding 1 to facilitate subsequent processing. Please see supplementary materials for more details.

B. Phase 2: Target Candidates Detection Based on Local Contrast and Gradient

Since the small targets typically have a higher intensity than their surrounding areas and are blurry due to atmospheric turbulence and the point spread function (PSF) of optical

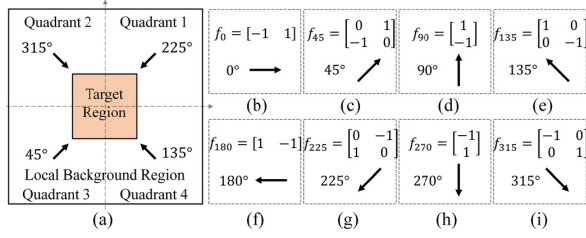


Fig. 2. (a) Illustration for the target and background region for candidate pixel and (b)–(i) gradient filters with different directions.

lens, we apply pixel-wise local contrast and gradient analysis simultaneously for target candidates represented by the mask generated in Phase 1, and the operation region is shown in Fig. 2(a). The size of the target and background region is $s \times s$ and $3s \times 3s$, respectively. The pixels in the same connected region share the same scale factor s , which is calculated in Phase 1.

Inspired by the calculation of local SCR [9], we proposed to generate a contrast map using

$$I_C(i, j) = \frac{(\mu_t - \mu_b)^2}{\sigma_b} \quad (3)$$

where (i, j) denotes the candidate pixel location, μ_t and μ_b are the mean intensity of candidate pixels in the target and background region in I_D , respectively, and σ_b is the standard deviation of candidate pixels in the background region. Since we apply the contrast analysis to the DoG filter outputs, it is obvious that the target region with a higher intensity than the local background region will be enhanced dramatically.

Since the noise and strong background clutter such as the cloud are also enhanced by the contrast map, we use the pixel-wise gradient feature, which can effectively distinguish the target from strong clutter in the input image I_{in} . We divide the region shown in Fig. 2(a) into four quadrants and perform gradient analysis for each quadrant separately.

According to the aforementioned characteristics of the IR target, the gradient distribution of local background regions points to the center of the target in most cases [5]. We obtain the eight directional gradient features for each quadrant by applying different filters, which are shown in Fig. 2(b)–(i). We only perform gradient filtering on four directions, including 0° , 45° , 90° , and 135° , and the result of the other four directions including 180° , 225° , 270° , and 315° can be simply obtained by inverting the results of 0° , 45° , 90° , and 135° , respectively. Let S_{ij} denote the sum of elements in matrix obtained by gradient filtering, where the value of i is 1–4 representing the four quadrants and the value of j is 1–8 representing 0° – 315° , respectively. The gradient score of each quadrant is given as follows:

$$G_{Q1} = 0.5 \times S_{16} + 0.25 \times S_{15} + 0.25 \times S_{17} \quad (4)$$

$$G_{Q2} = 0.5 \times S_{28} + 0.25 \times S_{21} + 0.25 \times S_{27} \quad (5)$$

$$G_{Q3} = 0.5 \times S_{32} + 0.25 \times S_{31} + 0.25 \times S_{33} \quad (6)$$

$$G_{Q4} = 0.5 \times S_{44} + 0.25 \times S_{43} + 0.25 \times S_{45}. \quad (7)$$

We assign higher weights to S_{16} , S_{28} , S_{32} , and S_{44} because 225° , 315° , 45° , and 135° denote the main directions of quadrants 1–4, respectively, as shown in Fig. 2(a). The intensity of

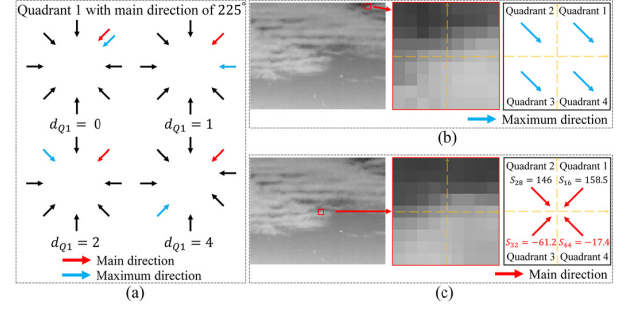


Fig. 3. (a) Demonstration of distance calculation for quadrant 1 in different situations, (b) example of the four quadrants having the same maximum directions, and (c) example of two quadrants obtaining negative scores in the main direction.

each candidate pixel (i, j) in the gradient map is defined by

$$I_G(i, j) = \frac{\mu_{GQ}}{(1 + p_{f1})} \times p_{f2} \quad (8)$$

where μ_{GQ} is the mean gradient score of the four quadrants, and both p_{f1} and p_{f2} are the penalty factors, which are used to suppress strong clutter. The penalty factor p_{f1} is formulated by

$$p_{f1} = \sum d_{Qi}, \quad i \in [1, 4] \quad (9)$$

where d_{Qi} represents the distance between the main direction index and the index yields the maximum gradient filtering result in each quadrant, and the demonstration of distance calculation for quadrant 1 in different situations is shown in Fig. 3(a).

The parameter p_{f2} is a binarized parameter, which is set to be 0 when the maximum directions in the four quadrants are the same or at least two quadrants obtain negative scores in the main direction because these two situations are highly related to the appearance of the strong edge, as shown in Fig. 3(b) and (c).

Finally, the detection score (map) I_S for the target candidate pixels is acquired by simply multiplying I_C and I_G to yield

$$I_S = I_C \times I_G. \quad (10)$$

The product of local contrast and gradient map can further enhance the target and suppress the clutter noise, as shown in Phase 2 of Fig. 1.

C. Phase 3: Target Detection via Iterative Segmentation

In Phase 3, instead of using a fixed threshold, we apply iterative segmentation to the target detection score I_S . The key idea of the iterative segmentation, which has the key advantage in distinguishing multiple targets with huge differences in detection scores, is to determine those pixels that meet the requirements in the descending order of detection score. The pseudo-code of iterative segmentation is described in Algorithm 2, where vectors V_S and V_I represent the intensity and indices for nonzero pixels in I_S , respectively. The parameters α and β are the customer-selected parameters ranging from 2 to 5. The parameter α depends on the detection score difference between the real target and clutter noise. The parameter β plays a significant role in balancing false and miss detection. The setting of the two parameters is given in the caption of Fig. 4.

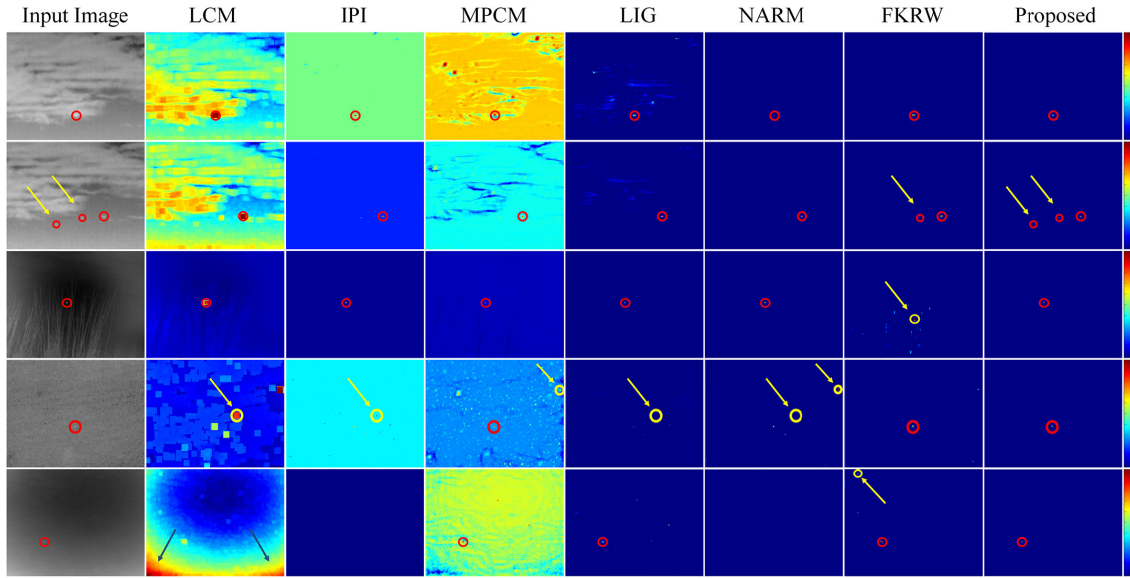


Fig. 4. Comparison of detection scores among different methods. The correct and false segmentation results are marked by red and yellow circles, respectively. The input images are the fourth frame in S1, the 28th frame in S1, the 64th frame in S2, the 20th frame in S3, and the 95th frame in S4, respectively. The segmentation parameters of the proposed method are $\alpha = 3$ and $\beta = 3$ for S1, $\alpha = 3$ and $\beta = 5$ for S2, $\alpha = 3$ and $\beta = 4$ for S3, and $\alpha = 3$ and $\beta = 5$ for S4, respectively.

Algorithm 2 Iterative Segmentation

Input: Detection score I_S .

Output: Binarized segmentation result I_{out} .

```

1: Generate vector  $V_S$  and  $V_I$ ,  $\mu_S = \text{mean}(V_S)$ .
2: while  $\text{flag} == 1$  do
3:    $V_{max} = \max(V_S)$ .
3:    $\mu = \text{mean}(V_S(\text{find}(V_S > 0)))$ ,
    $\sigma = \text{std}(V_S(\text{find}(V_S > 0)))$ .
4:   if  $V_{max} \geq \alpha \times \mu_S$  do
5:     return the index to  $I_{out}$ , clear the corresponding
       element in  $V_S$ .
6:   else if  $V_{max} \geq \mu + \beta \times \sigma$  do
7:     return the index to  $I_{out}$ , clear the corresponding
       element in  $V_S$ .
8:   else
9:      $\text{flag} = 0$ .
10:  end if
11: end while

```

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Comparison of Detection Performances

To compare the detection performances of different methods, we compared with LCM [7], IPI [9], MPCM [8], LIG [14], NRAM [10], and FKRW [12] using four IR sequences. The details of the test sequences are described in Table I. All the methods were performed by running them on a laptop with Intel i7-8665 and 16-GB RAM and with MATLAB 2019b.

The detection scores/maps of the proposed FAMSIS method and comparison methods are illustrated in Fig. 4. It should be noted that sequence one (S1) is a quite challenging data set because there are three targets in the last seven frames, two of which are quite dim and small. The proposed method has a clear advantage in detecting challenging targets such as missiles in the second row of Fig. 4. More importantly,

TABLE I
DETAILS OF THE TEST SEQUENCES

No.	Frame number	Target number	Resolution	Scene	Target size	Clutter description
S1	30	44	200x256	Cloudy sky	2x2~3x5	Bright target, strong clutter with irregular shape
S2	185	185	213x252	Grass and	2x3	Bright target with strong clutter in the surrounding
S3	50	50	152x208	sky	3x4	Quite dim target with noisy background and bad pixel
S4	200	200	208x256	Cloudy sky	3x3	Dim target in the edge of cloud

TABLE II
COMPUTATION TIMES COMPARISON USING DIFFERENT METHODS

		LCM	IPI	MPCM	LIG	NRAM	FKRW	Proposed
Running time (ms)	S1	26.7	4580.2	36.9	856.4	1123.4	46.7	16.7
	S2	25.6	4490.7	38.4	890.4	1382.4	88.7	11.8
	S3	22.4	1080.2	27.6	523.3	657.7	31.9	9.3
	S4	26.9	2593.5	33.3	936.3	1546.0	41.6	10.6

the proposed method performed most robustly for the dim and small targets with strong clutter noise, as shown in the fourth and fifth rows of Fig. 4.

Since the precision-recall curves (PRCs) are suitable for evaluating the detection performance on imbalanced data sets [15], [16], we used the PRC and computational times in our comparative studies. The PRCs on different sequences are shown in Fig. 5. It is clear that our method performed robustly on S1, S3 and S4. The proposed method is observed to be high performing for small spot-like targets with sky background, and the slightly inferior performance on S2 due to grass may affect the gradient map.

The comparison of running time among different methods is shown in Table II. Our method runs at least two times faster than the comparable methods.

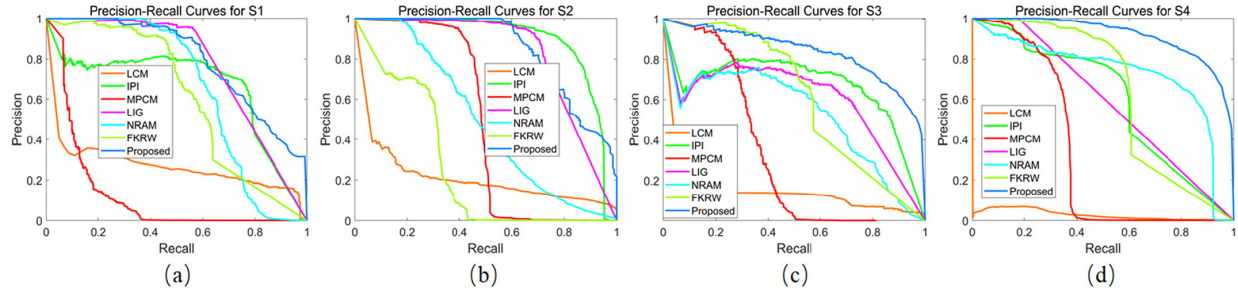


Fig. 5. PRCs of different methods. (a) PRC for S1. (b) PRC for S2. (c) PRC for S3. (d) PRC for S4.

TABLE III
METRICS COMPARISON USING DIFFERENT METHODS

	LCM	IPI	MPCM	LIG	NRAM	FKRW	Proposed
Detection rate (P_d)							
S1	68.18%	68.18%	65.91%	68.18%	65.91%	77.27%	90.91%
S2	98.92%	100%	100%	99.46%	100%	98.92%	99.46%
S3	88%	88%	96%	56%	94%	98%	98%
S4	100%	83.5%	96.5%	100%	79.5%	100%	100%
False alarm rate (P_f)							
S1	0.0%	9.09%	30.95%	6.25%	3.33%	8.11%	9.09%
S2	0.0%	0.54%	0.54%	6.12%	1.60%	4.69%	0.54%
S3	12%	18.52%	4%	41.67%	12.96%	3.92%	3.92%
S4	-	13.47%	3.5%	4.31%	0%	1.48%	0%
Averaged precision							
S1	0.2708	0.6621	0.1195	0.6538	0.7603	0.6234	0.7931
S2	0.2139	0.8895	0.4744	0.4944	0.8188	0.2540	0.8447
S3	0.1423	0.6711	0.3299	0.5256	0.6000	0.6453	0.8336
S4	0.0053	0.5961	0.3296	0.7287	0.5890	0.6733	0.9159

B. Comparison of Segmentation Performance

To evaluate the performance of our proposed iterative segmentation method, we also compared the binarized detection maps with other methods. The detection rate (P_d), false alarm rate (P_f), and the averaged precision (AP), which is the area under the PRCs computed from the segmented images, are shown in Table III. It can be seen that our proposed algorithm yielded the highest AP in three out of four cases. For S2, our AP is close to the top performers. It should be noted that the clutter has higher intensity than the target in S4 after the processing of the LCM. Therefore, we could not get the false alarm rate. The definition of the detection rate and false alarm rate is given as follows:

$$P_d = \frac{\text{Quantity of true targets detected in images}}{\text{Quantity of true targets existing in images}} \times 100\% \quad (11)$$

$$P_f = \frac{\text{Quantity of false targets detected in images}}{\text{Quantity of targets detected in images}} \times 100\%. \quad (12)$$

IV. CONCLUSION

In this letter, we proposed a novel and real-time method, which explicitly incorporates target size, local contrast, and gradient information to detect small IR targets. Moreover, the proposed method is flexible, modular, and high-performing. Extensive experimental results using four challenging IR image sequences clearly demonstrated that our proposed method performs robustly and runs at least two times faster

than other state-of-the-art methods, meaning that it will be promising in real-time applications.

Potential future directions include further improving the detection performance by combining the proposed method with multiframe-based methods and accelerate the processing procedure using GPU.

REFERENCES

- [1] X. Bai and Y. Bi, "Derivative entropy-based contrast measure for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2452–2466, Apr. 2018.
- [2] X. Guan, Z. Peng, S. Huang, and Y. Chen, "Gaussian scale-space enhanced local contrast measure for small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 327–331, Feb. 2020.
- [3] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," *Signal Data Process. Small Targets*, vol. 1954, pp. 2–11, May 1993.
- [4] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Signal Data Process. Small Targets*, vol. 3809, pp. 74–83, Oct. 1999.
- [5] S. Kim, Y. Yang, J. Lee, and Y. Park, "Small target detection utilizing robust methods of the human visual system forIRST," *J. Infr., Millim., THz Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.
- [6] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infr. Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.
- [7] C. L. Philip Chen, H. Li, Y. Wei, T. Xia, and Y. Yan Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
- [8] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, Oct. 2016.
- [9] 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.
- [10] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint l2, l1 norm," *Remote Sens.*, vol. 10, no. 11, pp. 1–26, 2018.
- [11] H. Zhang, L. Zhang, D. Yuan, and H. Chen, "Infrared small target detection based on local intensity and gradient properties," *Infr. Phys. Technol.*, vol. 89, pp. 88–96, Mar. 2018.
- [12] Y. Qin, L. Bruzzone, C. Gao, and B. Li, "Infrared small target detection based on facet kernel and random walker," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7104–7118, Sep. 2019.
- [13] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, p. 382, Feb. 2019.
- [14] W. Davidson and M. Abramowitz, *Molecular Expressions Microscopy Primer: Digital Image Processing Difference of Gaussians Edge Enhancement Algorithm*. Tallahassee, FL, USA: Olympus America Inc. and Florida State Univ., 2006.
- [15] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, pp. 1–21, Mar. 2015.
- [16] K. A. Islam, M. S. Uddin, C. Kwan, and J. Li, "Flood detection using multi-modal and multi-temporal images: A comparative study," *Remote Sens.*, vol. 12, no. 15, p. 2455, Jul. 2020.