

福昕PDF编辑器

• 永久 • 轻巧 • 自由

升级会员

批量购买



永久使用

无限制使用次数



极速轻巧

超低资源占用，告别卡顿慢



自由编辑

享受Word一样的编辑自由



手机 扫一扫，关注公众号

A Double-Neighborhood Gradient Method for Infrared Small Target Detection

Lang Wu, Yong Ma^{ID}, Fan Fan^{ID}, Minghui Wu, and Jun Huang^{ID}, *Member, IEEE*

Abstract—Effective and efficient infrared (IR) small target detection is essential for IR search and tracking (IRST) systems. The current methods have some limitations in background suppression or detection of targets close to each other. In this letter, a double-neighborhood gradient method (DNGM) is proposed. First, a new technology of the tri-layer sliding window is designed to measure the double-neighborhood gradient. Then, the DNGM is obtained by multiplying the double-neighborhood gradient. In this way, even the sizes of the targets may vary, ranging from 2×1 to 9×9 pixels, the target can be better highlighted under a fixed scale, and background interference can be suppressed. Finally, the target is segmented from the DNGM salience map by an adaptive threshold. Experiments illustrate that the proposed method can avoid the “expansion effect” of the traditional multiscale human vision system (HVS) method and can accurately detect multiple targets close to each other. Besides, the proposed method is more robust and real-time than the existing methods.

Index Terms—Double-neighborhood gradient, infrared (IR) small target, non multiscale, tri-layer window.

I. INTRODUCTION

AS ONE of the core parts of many military systems such as early warning and precision guidance, infrared (IR) small target detection plays an important role and has a strong impact on system performance [1], [2]. IR small targets usually contain a few to dozens of pixels in each frame, due to remote imaging. The lack of obvious texture features and shape information [3], as well as the complex background and noise, brings great difficulties to small target detection and tracking. Typical IR small target detection methods can be roughly categorized into track-before-detection (TBD) method and detection-before-track (DBT) method. At the same time, there are some methods that combine these two kinds of methods, such as [4], in which the researchers build a spatiotemporal saliency model by combining single-frame detection with track prediction. The DBT method is widely

Manuscript received May 20, 2020; revised June 11, 2020; accepted June 13, 2020. This work was supported in part by the Natural Science Foundation of Hubei Province under Grant 2019CFCB162 and Grant 2018CFA006 and in part by the Fundamental Research Funds for the Central Universities under Grant 2042020kf0017. (*Corresponding author: Jun Huang.*)

Lang Wu and Minghui Wu are with the Electronic Information School, Wuhan University, Wuhan 430072, China (e-mail: wulang@whu.edu.cn; wmhcrystal@whu.edu.cn).

Yong Ma, Fan Fan, and Jun Huang are with the Electronic Information School, Wuhan University, Wuhan 430072, China, and also with the Institute of Aerospace Science and Technology, Wuhan University, Wuhan 430079, China (e-mail: mayong@whu.edu.cn; fanfan@whu.edu.cn; junhwong@whu.edu.cn).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2020.3003267

used in practical projects due to its low complexity and easy hardware implementation. Thus, in this letter, we will focus on the DBT method.

Early DBT methods such as Top-hat [5], infrared patch-image model (IPI) [6], and DSVGCI [7] generally have the problem of poor performance in detection with complex background. Many relatively new DBT methods have been proposed in the past few years. The methods based on low rank and sparsity [8], [9] have made good progress in single-target detection. However, they have limited ability to suppress the interference with higher intensity in complex background. The methods based on machine learning [10], [11] and deep learning depend on training samples and features, and hence in practical application, it is difficult to include all kinds of complex background interference in limited training data. The methods based on human visual system [12]–[15] methods have attracted much attention from relevant scholars because of their good detection and real-time performance. However, because traditional multiscale human vision system (HVS)-based methods do not consider the situation that multiple targets are close to each other, it is easy to identify multiple targets as one target, resulting in missed detection. The reason is that adopting multiscale technology will make the target of less than 9×9 pixels in raw image displayed as 9×9 pixels in the detection result [12], [13], [15]. We call it the “expansion effect” (see Section II-A). In this case, the detection probability of multiple targets is reduced, and the multiscale process is also time consumption.

For the sake of achieving better detection performance and solving the problem of the “expansion effect,” we propose a double-neighborhood gradient method (DNGM). First, we design a novel tri-layer sliding window to ensure the ability to detect multiscale targets under a fixed scale. Because the proposed method does not adopt multiscale technology, we can avoid the “expansion effect” and spend less time. Then a measure based on the double-neighborhood gradient is presented to improve the saliency of the target and suppress the background interference better. Finally, the experimental results show that the proposed method is more robust and real-time than several widely used methods and can accurately detect targets close to each other.

II. PROPOSED METHOD

The flowchart of the proposed method is presented in Fig. 1. First, the novel tri-layer window is designed to detect targets with different sizes under a fixed scale. Then the DNGM is calculated to enhance the target and suppress the background. Finally, the real target is extracted by threshold operation.

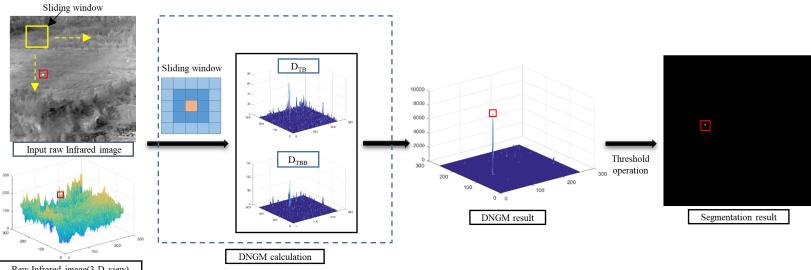


Fig. 1. Overview of the proposed method.

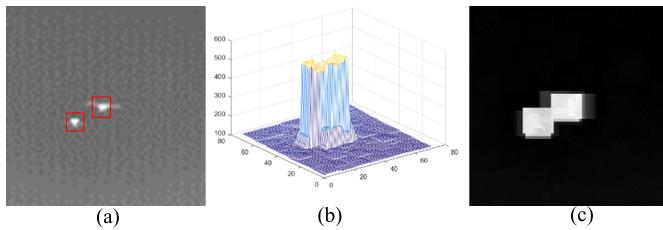


Fig. 2. Expansion effect of the multiscale method. (a) Infrared image with two targets. (b) Multiscale result (3-D view). (c) Two targets overlap.

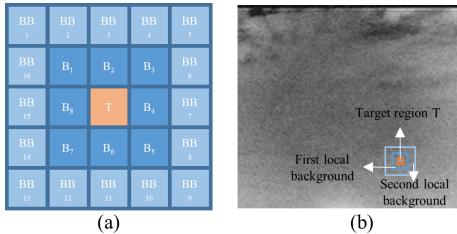


Fig. 3. Tri-layer window of the DNGM. (a) Nested structure of tri-layer window. (b) Regions of tri-layer window.

A. Construction of the Tri-Layer Window

Traditional local contrast measure (LCM) and improved LCM (ILCM), relative LCM (RLCM), and difference-based local contrast method (DLCM) adopt multiscale technology to detect small targets of 3×3 to 9×9 pixels. For small targets with a size less than 9×9 pixels in the IR image, the multiscale technology adopted by the LCM will enhance the background area around the target, resulting in the size of the detected target enlarging to 9×9 pixels. We call this the “expansion effect.” As shown in Fig. 2, the “expansion effect” makes two targets overlapped. So the quantity of the targets cannot be detected accurately. Because multiscale processing usually takes the maximum response value under multiple scales as the final output, as long as the size of the window cell is larger than the actual target size in the raw image, the background close to the target will be enhanced as well, and the detected target size will be expanded. Therefore, the key of dealing with this effect is to find a non multiscale way that can detect targets of different sizes adaptively.

To detect small targets from 2×1 to 9×9 pixels under a fixed scale, we design a new sliding window by adding a layer of peripheral neighborhood cells BB_1-BB_{16} based on the window of LCM with cell size of $N \times N$ pixels. As shown in Fig. 3, the entire window is divided into three regions, where cell T is the reference cell (representing the region where the target could appear). B_i around T is the first neighborhood background with a total of eight cells. The outermost BB_j is

the second neighborhood background with 16 cells. When the target locates in the center of a tri-layer window, the grayscale difference will be reflected in the three regions regardless of the size of the small target, and these differences can be used to detect targets of different sizes.

B. DNGM Calculation

The purpose of small target detection is to enhance the target contrast as much as possible to highlight the small targets effectively. According to our tri-layer window, we can make full use of the differences between the three regions to measure contrast. The minimum gray gradient between the center cell and the neighborhood cell can be used as a definition of contrast. Therefore, the difference between the center cell T and the neighborhood cell can be defined as

$$D_{TB} = \begin{cases} m_T - \max(m_{B_i}), & m_T > \max(m_{B_i}) \\ 0, & \text{else} \end{cases} \quad (1)$$

$$D_{TBB} = \begin{cases} m_T - \max(m_{BB_j}), & m_T > \max(m_{BB_j}) \\ 0, & \text{else.} \end{cases} \quad (2)$$

D_{TB} and D_{TBB} represent the least gradient between the reference cell and the first and the second neighborhood cells, respectively. Herein, m_T , m_{B_i} , and m_{BB_j} are the mean value of the reference cell, the i th first-neighborhood background cell, and the j th second-neighborhood background cell, respectively. Use the mean grayscale difference to calculate the gradient, which can effectively diminish the impact of the highlighted background.

In addition, we enhance the saliency of small targets by multiplying the double-neighborhood gradient, which can effectively use the two neighborhood gradient. In this way, the region with the largest gradient product is the most likely to be the target region. It is possible to exclude the situation where one of the D_{TB} and D_{TBB} is larger and the other is zero, which reduces the false alarm rate of our method. Finally, based on the central pixel of the reference cell, we give DNGM which is obtained as follows:

$$\text{DNGM} = D_{TB} * D_{TBB}. \quad (3)$$

When our window scans background and small targets with different sizes, the following situations may occur.

- 1) When our window moves to a target of a smaller size [see Fig. 4(a)], D_{TB} and D_{TBB} of the center pixel are usually larger, so the corresponding DNGM is very large, and the target is enhanced.

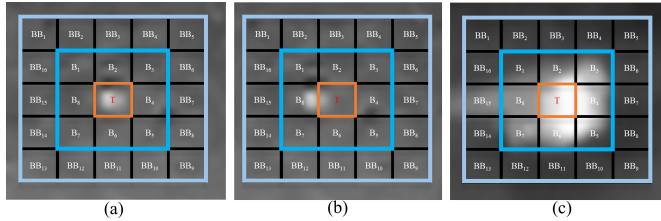


Fig. 4. Illustration of the tri-layer window sliding on an image. (a) Window detects the smaller target. (b) Window detects the neighborhood of the smaller pixels' target. (c) Window detects the larger target.

- 2) When our window moves to the background area around the small target [see Fig. 4(b)], D_{TB} is equal to 0, and D_{TBB} is small. According to (3), we can easily get that DNGM value is equal to 0, which means the neighborhood background of the target is suppressed. Therefore, the proposed method does not have the “expansion effect” on small-size targets.
- 3) When our window moves to a small target of a larger size [see Fig. 4(c)], considering that the middle part of the target is slightly convex, D_{TB} is usually greater than 0. At the same time, although the gray value inside the small target is not the maximum value of the entire image, it is usually greater than the gray value of a small area around the target, so D_{TBB} is larger. In the end, DNGM is larger, and the large-size target is also enhanced.
- 4) Because the minimum gradient is used between the cell T and the background cells, when the window is on the background or the highlighted edge, we can easily get that D_{TB} and D_{TBB} are approximately equal to 0, and the DNGM approaches 0.

After the DNGM calculation, different types of background interference will be suppressed, and the real small target is significant.

C. Target Extraction

In the light of the calculation of DNGM, we can get the DNGM saliency map of each IR image. Considering that the saliency map has different levels of clutter background, we segment the target and background by adopting the adaptive thresholding method. In this letter, the adaptive threshold R is calculated by

$$R = \mu_{\text{DNGM}} + k\delta_{\text{DNGM}} \quad (4)$$

where μ_{DNGM} and δ_{DNGM} are the mean value and the standard deviation (STD) of the DNGM map, respectively, and k is a given parameter. The specific values of k will be tested in Section III.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we use both the simulation images and real IR image sequences to verify the effectiveness and robustness of the proposed method. In addition, in this letter, the parameter k of our method in the experiments is set to be 40, and the parameters of the comparison methods are set to the recommended values in their articles. Seq. 1 [16] is a multitarget real IR image sequence, and Seq. 2 and

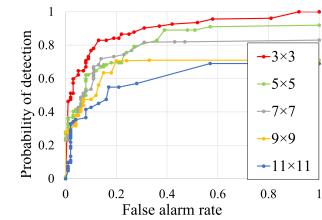


Fig. 5. ROC curves of Seq. 1 for different cell sizes.

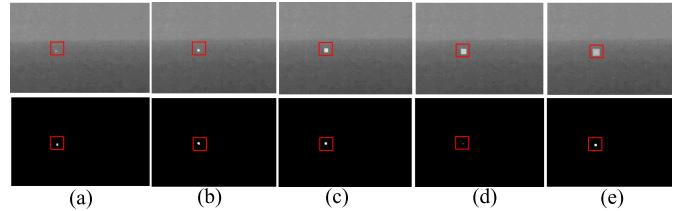


Fig. 6. Detection effect of our method on targets of different sizes. (a) 1 × 2. (b) 3 × 3. (c) 5 × 5. (d) 7 × 7. (e) 9 × 9

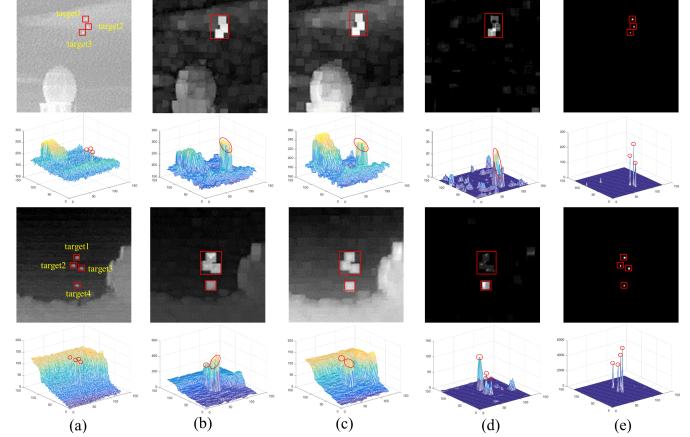


Fig. 7. Raw images and detection results of multitarget simulation images. (a) Simulation images. (b) LCM results. (c) ILCM results. (d) RLCM results. (e) Results of our method.

Seq. 3 [17] are single-target real IR image data sets with complex backgrounds. The experiments were conducted in MATLAB 2016b on a laptop with 16-GB RAM and 2.10-GHz AMD R5 CPU.

A. Evaluation Metrics and Comparison Methods

To assess the performance of different IR small target detection methods quantitatively, the signal-to-clutter ratio gain (SCRG), background suppression factor (BSF), receiver operating characteristic (ROC) curve, and time consumption are adopted [18].

The vertical coordinate of the ROC curves is the detection probability P_d and the horizontal coordinate is the false alarm rate P_f which are defined in the following equations:

$$P_d = \frac{\text{number of true targets detected}}{\text{number of actual targets}} \quad (5)$$

$$P_f = \frac{\text{number of false detections}}{\text{number of tested frames}}. \quad (6)$$

In this letter, Top-hat [5], LCM [12], ILCM [13], RLCM [15], multiscale patch-based contrast measure (MPCM) [14], IPI [6], average difference weighted by cumulative directional derivatives (AACDD) [19], nonconvex rank approximation

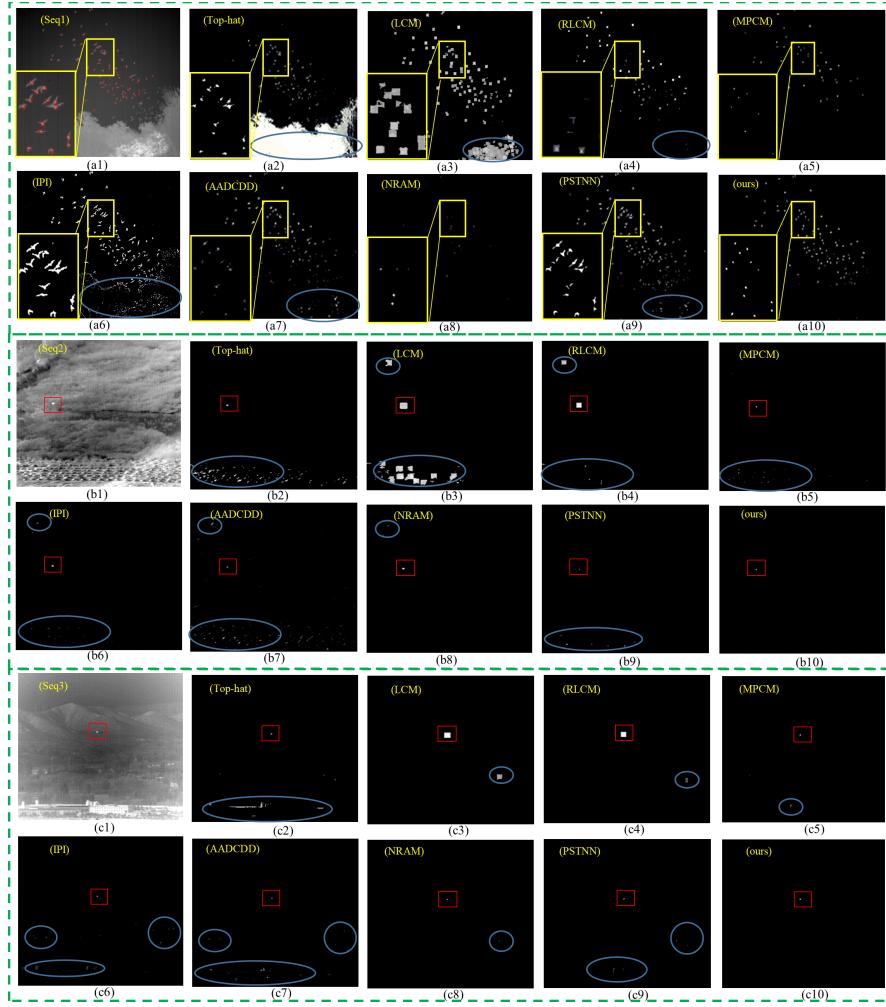


Fig. 8. Raw images and experimental results of different real IR image sequences.

minimization (NRAM) [8], and partial sum of tensor nuclear norm (PSTNN) [9] are chosen as the comparison methods.

B. Experimental Results and Analysis

Because the proposed method adopts a tri-layer window, it is possible to detect targets of different sizes under a fixed cell size theoretically. To verify the effect of different window cell sizes on the performance of the proposed method, we use Seq. 1 to test different cell sizes, and the ROC curves are shown in Fig. 5. It can be easily seen that 3×3 pixels is the best value. Therefore, in the following experiments, our window cell size is fixed to 3×3 pixels.

The proposed method is tested with simulation targets of size 1×2 , 3×3 , 5×5 , 7×7 , and 9×9 pixels. The test result is presented in Fig. 6, which shows that our method can achieve the detection of targets with different sizes under a fixed scale of 3×3 pixels. Meanwhile, the proposed method and three multiscale HVS methods are tested with multitarget simulation images, and the result is shown in Fig. 7. We find that the proposed method successfully overcomes the “expansion effect,” and it can accurately detect the quantity and the location of multiple small targets. LCM, ILCM, and RLCM enlarge the target due to their multiscale technology, resulting in multiple targets in the saliency map sticking

together or even overlapping. Besides, in their results, the quantity of the targets detected is less than the actual quantity.

In Fig. 8, the final detection outcome of nine methods for the real IR image sequences is given. We can get that the method we presented can detect multiple targets that are close to each other well and truly and make the targets salient while various complex background interferences are effectively suppressed. In the results of Seq. 1, the LCM saliency maps have overlapping targets. Top-hat, RLCM, and NRAM results have many targets undetected, and NRAM detected only a few targets, far less than the quantity of the actual target. IPI, AACDD, and PSTNN show good detection results, and IPI and PSTNN retain some shape information of the target. In the detection results of Seq. 2 and Seq. 3, a lot of false alarms emerge when the comparison methods detected the targets. Top-hat and LCM cannot suppress complex backgrounds and generate a large number of false alarms or even miss targets in all sequences. RLCM generates residual clutter on the bright speckle background in Seq. 1 and Seq. 2. Although MPCM can get good detection results by increasing the threshold value, it is easy to identify the local dark background as the false alarm. IPI has a poor suppression effect on the highlighted edge. PSTNN cannot effectively suppress the top corner of the highlighted background. For example, PSTNN generates

TABLE I
SCRG, BSF, AND TIME CONSUMPTION OF NINE METHODS

Evaluation Indicators	SEQUENCES	Top-hat	LCM	RLCM	MPCM	IPI	AADCDD	NRAM	PSTNN	Ours
SCRG	1	3.180	1.081	4.238	7.102	0.972	8.770	137.618	10.983	37.626
	2	4.171	0.570	0.574	14.688	1.425	22.350	Inf	128.908	Inf
	3	3.866	0.562	0.558	10.419	1.247	1.166	Inf	50.481	Inf
BSF	1	3.399	1.050	13.831	14.262	4.304	32.552	159.392	18.717	55.770
	2	4.422	3.370	11.238	24.954	21.642	24.835	108.205	10.498	121.283
	3	9.808	5.294	14.980	44.440	45.953	40.212	121.769	29.490	141.077
Time(s)	1	0.038	0.284	48.838	0.332	719.894	0.149	135.833	136.231	0.178
	2	0.011	0.084	11.802	0.083	12.478	0.036	5.337	0.420	0.045
	3	0.011	0.082	11.716	0.084	12.048	0.035	6.933	0.237	0.047

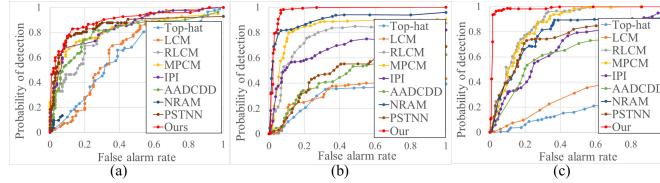


Fig. 9. ROC curves of three sequences (a) Seq.1. (b) Seq.2. (c) Seq.3.

a noticeable false alarm at the top corner of a building at the bottom of the image in Seq. 3.

Table I shows the average SCRG, BSF, and time consumption of different methods for three sequences. It is worth noting that Inf (i.e., infinity) means that the clutter in the area around the target is fully removed. Table I shows that our method has stronger robustness and runs faster in comparison to others. We can get that the SCRG and BSF of the NRAM and the proposed method in Seq. 1 far outweigh other methods, but NRAM only detects a few targets and misses a large number of real targets. In both Seq. 2 and Seq. 3, our method has the largest SCRG and BSF. In terms of average time consumption, among the eight comparison methods, only Top-hat and AADCDD run faster than our method.

Fig. 9 demonstrates the ROC results of nine methods on three sequences. It can be drawn that our method has a great advantage over other methods and is more suitable for the detection of close multiple targets.

IV. CONCLUSION

In this letter, a DNGM is proposed. The main idea of the DNGM is to design a tri-layer sliding window to detect small targets of different sizes under a fixed scale so that the “expansion effect” caused by multiscale technology is avoided and the computation complexity is much lower. Moreover, the saliency map obtained by multiplying the double-neighborhood gradient can make the target salient better while suppressing the background interference. The experimental results have demonstrated that the proposed method has strong robustness and fast detection speed and has greater advantages, especially for detecting targets close to each other, in comparison to other methods.

REFERENCES

- [1] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, “Max-mean and max-median filters for detection of small targets,” *Proc. SPIE*, vol. 3809, pp. 74–83, Oct. 1999.
- [2] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, “FusionGAN: A generative adversarial network for infrared and visible image fusion,” *Inf. Fusion*, vol. 48, pp. 11–26, Aug. 2019.
- [3] J. Ma, Y. Ma, and C. Li, “Infrared and visible image fusion methods and applications: A survey,” *Inf. Fusion*, vol. 45, pp. 153–178, Jan. 2019.
- [4] Y. Li, Y. Zhang, J.-G. Yu, Y. Tan, J. Tian, and J. Ma, “A novel spatio-temporal saliency approach for robust dim moving target detection from airborne infrared image sequences,” *Inf. Sci.*, vol. 369, pp. 548–563, Nov. 2016.
- [5] M. Zeng, J. Li, and Z. Peng, “The design of top-hat morphological filter and application to infrared target detection,” *Infr. Phys. Technol.*, vol. 48, no. 1, pp. 67–76, Apr. 2006.
- [6] 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.
- [7] C. Yang, J. Ma, S. Qi, J. Tian, S. Zheng, and X. Tian, “Directional support value of Gaussian transformation for infrared small target detection,” *Appl. Opt.*, vol. 54, no. 9, pp. 2255–2265, Mar. 2015.
- [8] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, “Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm,” *Remote Sens.*, vol. 10, no. 11, p. 1821, Nov. 2018.
- [9] 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.
- [10] Y. Bi, X. Bai, T. Jin, and S. Guo, “Multiple feature analysis for infrared small target detection,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1333–1337, Aug. 2017.
- [11] J. Ryu and S. Kim, “Heterogeneous gray-temperature fusion-based deep learning architecture for far infrared small target detection,” *J. Sensors*, vol. 2019, pp. 1–15, Aug. 2019.
- [12] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, “A local contrast method for small infrared target detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] Z. Wu, N. Fuller, D. Theriault, and M. Betke, “A thermal infrared video benchmark for visual analysis,” in *Proc. 10th IEEE Workshop Perception Beyond Visible Spectrum (PBVS)*, Columbus, OH, USA, 2014, pp. 201–208. [Online]. Available: <http://vcipl-okstate.org/pbvs/bench/index.html>
- [17] A Dataset for Dim-Small Target Detection and Tracking of Aircraft in Infrared Image Sequences. Accessed: Jun. 23, 2020. [Online]. Available: <http://www.csdata.org/p/3871/>
- [18] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, “Infrared small-dim target detection based on Markov random field guided noise modeling,” *Pattern Recognit.*, vol. 76, pp. 463–475, Apr. 2018.
- [19] S. Aghaziyarati, S. Moradi, and H. Talebi, “Small infrared target detection using absolute average difference weighted by cumulative directional derivatives,” *Infr. Phys. Technol.*, vol. 101, pp. 78–87, Sep. 2019.