



Small Infrared Aerial Target Detection Using Spatial and Temporal Cues

Liangchao Guo, Wenlong Zhang, Xiaoliang Sun^(✉), Zi Wang, and Yang Shang

National University of Defense Technology, Changsha, China
alexander_sxl@nudt.edu.cn

Abstract. Due to rapidly relative motion between target and infrared imaging platform, clutter background, etc., robust small infrared aerial target detection is still an open problem. A novel small infrared aerial target detection method using spatial and temporal cues is proposed in this paper. First, using spatial cues, we take target candidate detection as a binary classification problem. Target candidates in each single frame are detected via interesting pixel detection and a trained LightGBM model. Then, using temporal cues, we model the local smoothness and global continuous characteristic of the target trajectory as short-strict and long-loose constraints. The trajectory constraints within image sequence are used in detecting the true small infrared aerial targets from numerical target candidates. Experiment results on the public dataset SIATD show that the proposed method performs better than other existing methods in detecting small infrared aerial target and shows great robustness toward clutter backgrounds.

Keywords: Aerial · Small infrared target detection · Spatial and temporal cues · Trajectory constraint · LightGBM

1 Introduction

Infrared detection has advantages of all-day and all-weather operation, high resolution, etc. Thus, infrared detection system has been widely used in air platform early warning, guidance, etc. [1]. Due to the rapidly relative motion and clutter backgrounds, small aerial target detection is still an open problem for the airborne infrared detection system.

Existing infrared detection methods can be categorized into single frame based methods and multiple successive frames based methods. Single frame based methods detect the small infrared target mainly using the differences between the target and background. The small infrared target is often modeled as a spot target of isotropic distribution and this kind method is easy to be implemented and efficiently. However, cues provided by a single frame may be inadequate for robust small infrared target detection. The temporal cues contained in multiple successive frames are important to robust small infrared target detection [1]. Multiple successive frames based methods boost the performance of small target detection by associating multiple image data. However, the adoption of the temporal cues increases the computational complexity. And such existing methods have

trouble in detecting small aerial targets for airborne infrared detecting systems focused in this paper.

Existing works have demonstrated that temporal cues play an important role in robust small infrared target detection, especially for complex cases. For the targets, we find that the true targets exhibit continuous and smooth long trajectories while the clutter does not. Based on the facts, this paper tackles the problem of small infrared aerial target detection by using spatial and temporal cues. The proposed method firstly detects the target candidates from each frame using interesting pixel selection and a trained highly efficient gradient boosting decision tree (LightGBM) [2] model. Then based on the short-strict and long-loose constraints, the true targets are detected from numerical target candidates by trajectory segment growth and merging. Experimental results indicate that our method can detect the small infrared aerial targets robustly and achieves superior performance than other existing methods.

2 Related Works

Most related works to this paper are summarized in this section. As mentioned above, existing small infrared target detection methods can be classified into single frame and multiple frames based methods.

2.1 Single Frame Based Methods

Single frame based methods detect the small infrared targets from a single image. They are mainly using the difference between the target and the background.

Moradi et al. [3] modeled the spot target using the point spread function of the imaging system. Such a method is simple and efficient. However, it performs poorly under clutter backgrounds. Based on the remote imaging and infrared imaging characteristics, researchers modeled the background as approximately common or uniform components. Compared to the background, the small target has a small spatial spread. The small target is detected by subtracting the estimated background. Gao et al. [4] modeled the background using Infrared Patch-Image (IPI) reconstruction. Xue et al. [5] introduced multiple sparse constraints in reconstruction. These methods usually work well only when the background satisfies the assumption of large space expansion, and the background modeling takes a lot of time. Zhao et al. [6] extract spatial size and contrast information of small infrared targets in the max-tree and min-tree and proposed a novel detection method based on multiple morphological profiles. The method suffers a high false alarm rate for clutter backgrounds.

For CNN-related methods, Dai et al. [7] preserved and highlighted the small target feature by exploiting a bottom-up attentional modulation integrating the low-level features into the high-level features of deeper layers. And the authors of [8] constructed a Generative Adversarial Networks (GAN), upon U-Net, to learn the features of small infrared targets and directly predict the intensity of targets.

2.2 Multiple Frames Abased Methods

Single frame based methods have trouble in detecting small infrared targets under clutter backgrounds. Multiple frames based methods introduce temporal cues in target detecting via association between successive frames.

Marco et al. [9] proposed a generalized likelihood ratio test based method for small target detection in sea background. [10] adopted passion distribution in energy accumulation for small infrared target detection. Single pixel association methods are sensitive to clutter backgrounds or isolated spot noise. Li et al. [11] enhanced the small infrared target via saliency analysis based on motion and appearance. The spatio-temporal tensor model is adopted to model the background in [12, 13]. The multiple subspace learning is adopted to modify [14] in [15]. Such methods cannot handle rapidly changing backgrounds well. In addition, the methods are often complex and cannot meet the needs of real-time applications.

Most of existing exploration works are single frame based methods. Temporal cues have not been adopted in the mentioned works. Although the performance of related works has been improved, they have difficulties in detecting small infrared aerial targets under clutter backgrounds.

3 Small Infrared Aerial Target Detection using Spatial and Temporal Cues

The proposed method firstly detects target candidates from each frame using a trained LightGBM model. Then the trajectory candidates are generated by linking the target candidates within successive frames. The true trajectories that meet the short-strict and long-loose constraints are detected finally.

3.1 Target Candidate Detection for Each Frame

Due to remote imaging, the small aerial targets are presented as spot targets in the image. And, the spot targets may brighter or darker than their surroundings. This paper takes the target candidate detection for each frame as a binary classification problem. For each interesting pixel, we extract features in the local region centered on it. Then the trained LightGBM model takes the features as input and determines the pixel is a target candidate or not.

Interesting Pixel Detection. The small infrared targets only correspond to a small part of pixels in the images. In order to detect the small infrared target efficiently, we extract the interesting pixels which are more likely to be the target from the image and enter the following process. As mentioned above, the small infrared aerial target focused in this paper is often presented as a spot target. It may be brighter or darker than its neighbors. This paper adopts positive and negative median filter to detect interesting pixels, as shown in Eq. (1).

$$Label(x, y) = \begin{cases} 1, & I(x, y) > (k_1 + median((x, y))) \vee I(x, y) < (median((x, y)) + k_2) \\ 0, & otherwise \end{cases} \quad (1)$$

For the input image I , $Label(x, y)=1$ indicates the pixel (x, y) is an interesting pixel and $Label(x, y)=0$ denotes not. $median((x, y))$ calculates the median value in a local region of certain size centered on (x, y) . We set the parameters $k_1 = 15$ to select the brighter targets and $k_2 = -20$ to select the darker targets.

Feature Extraction and Learning. This paper adopts LightGBM [2] in target candidate detection. This paper extracts features for each interesting pixel in the local region which is as rectangle centered on it and computes 7 features from the local region, including kurtosis γ_2 , skew S_k , entropy H , mean μ , variance σ^2 , maximum v_{\max} and minimum v_{\min} . Let $L_{R_1 \times R_2}(x, y)$ denotes the local rectangle region centered on the pixel (x, y) of size $R_1 \times R_2$. We flat the region into vector $\mathbf{V}=\{v_0, v_1, \dots, v_{N-1}\}_{N=R_1 \times R_2}$. The definitions of the 7 features are as Eq. (2) shows. $p(\cdot)$ denotes the probability of the intensity value and it can be inferred from the intensity histogram of the input image.

$$\begin{aligned}
 \text{kurtosis : } \gamma_2 &= \frac{\mu_4}{\sigma^4} - 3 & \text{variance : } \sigma^2 &= \frac{1}{N} \sum_{v_i \in \mathbf{V}} (v_i - \mu) \\
 \text{skew : } S_k &= \frac{\mu_3}{\sigma^3} & \text{maximum : } v_{\max} &= \max_{v_i \in \mathbf{V}}(v_i) \\
 \text{entropy : } H(\mathbf{V}) &= - \sum_{v_i \in \mathbf{V}} p(v_i) \log p(v_i) & \text{minimum : } v_{\min} &= \min_{v_i \in \mathbf{V}}(v_i) \\
 \text{mean : } \mu &= \frac{1}{N} \sum_{v_i \in \mathbf{V}} v_i & &
 \end{aligned} \tag{2}$$

In the training dataset, the small infrared aerial targets are annotated. We take the pixels within the regions centered on the label positions of size 3×3 as positive samples. The remaining pixels in the images can be taken as negative samples. The 7 dimension feature vector by multi-scale processing strategy is calculated for each sample to train the LightGBM model.

3.2 Target Detection Using Trajectory Constraints

The target candidates in each frame have been detected as mentioned above. We use the homograph transform to model the inter-frame movement in this paper. The registration between successive frames is built by SURF [16] feature point extraction and matching. Then we remap the target candidates in each frame to the coordinate of the first frame within the time window.

We intend to detect the true targets whose trajectories obey the short-strict and long-loose constraints. In the captured image sequence, the true target form a continuous and smooth long trajectory. The long trajectory can be used to distinguish targets from back-grounds robustly. This paper models the target's movement as a piecewise uniform motion. We impose the strict constraint, uniform motion, on the trajectory in a small time interval to eliminate the interference of noise as much as possible. On the contrary, we impose the relax constraint on the trajectory in a long time range to extend the length of the trajectory as much as possible. The trajectory synthesizing and validation include trajectory segment growth and merging. They are detailed as the following.

Trajectory Segment Growth with Short-Strict Constraint. Trajectory segment growth links the target candidates in the current frame to the existing trajectory segments properly. This paper models the target movement as uniform motion, i.e. under the short-strict constraint, in a small time interval. We set the small time interval is 3 successive frames in this paper. Trajectory segments are growing under the short-strict constraint. Given the existing trajectory segment set $\{T^i\}_M$ and the target candidate set $\{c_j^t\}_N$ in the current frame t .

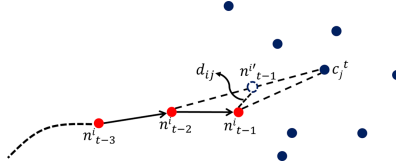


Fig. 1. Trajectory segment growth with short-strict constraint.

As shown in Fig. 1, we take a sample trajectory segment $T^i = \{\dots, n_{t-3}^i, n_{t-2}^i, n_{t-1}^i\}$ to detail the implementation of trajectory segment growth. n_{t-1}^i is the detected target in the last frame ($t-1$). Under the uniform motion constraint in the small time interval, we define the cost of linking c_j^t to T^i . The link involves n_{t-2}^i, n_{t-1}^i and c_j^t . Using n_{t-2}^i and c_j^t , we get the ideal middle point $n_{t-1}^{i'}$ under the uniform motion constraint. d_{ij} is the Euclidean distance between n_{t-1}^i and $n_{t-1}^{i'}$. The cost $C(i, j)$ of the link is defined as Eq. (3).

$$C(i, j) = \frac{d_{ij}}{\|n_{t-2}^i - n_{t-1}^i\|_2} = \frac{\|n_{t-2}^i + c_j^t - 2n_{t-1}^i\|_2}{2 \times \|n_{t-2}^i - n_{t-1}^i\|_2} \quad (3)$$

The smaller the $C(i, j)$ is, the more the link is met with the short-strict constraint. The cost matrix C contains all possible links' cost values. We define the binary linking matrix A_1 in Eq. (4).

$$A_1(i, j) = \begin{cases} 1, & C(i, j) \leq \sigma_1 \\ 0, & C(i, j) > \sigma_1 \end{cases} \quad (4)$$

σ_1 is the cost threshold and is set 0.2. We also restrict the absolute velocity value of the target. The restriction for the target candidates c_j^{t-1} and c_j^t in successive frames is defined as Eq. (5). σ_2 , a velocity threshold, is set 10.

$$CoV(c_j^{t-1}, c_j^t) = \begin{cases} 1, & \|n_{t-2}^i - n_{t-1}^i\|_2 \leq \sigma_2 \\ 0, & \|n_{t-2}^i - n_{t-1}^i\|_2 > \sigma_2 \end{cases} \quad (5)$$

For an existing trajectory segment, if more than one target candidate meet Eq. (3–4), we link the trajectory segment to each target candidate and record every new link as a new

trajectory segment. In order to find new target, we also link the target candidates $\{c_j^t\}_N$ in the current frame t and the candidates $\{c_j^{t-1}\}_{N'}$ in the last frame $(t-1)$ under Eq. (5) to form new trajectory segments.

Trajectory Segment Merging with Long-Loose Constraint. Due to noise interference or clutter background, it results in that the true target trajectory is divided into several segments. Trajectory segment merging intends to link the trajectory segments which correspond to the same target. The merging is performed according to the similarity between trajectory segments. Figure 2 presents three track segment pairs with different relative positions. Comparing to the segments in Fig. 2(a, c), the segments in Fig. 2(b) are more likely to corresponding to the same target.

This paper summarizes the features of the track segments which are corresponding to the same target as follows: (a) The track segments do not overlap in time. (b) The extension of the track segments is close to each other. (c) The velocity values of different track segments are close to each other. We take two track segments and as samples to detail the definition of the similarity measure, as shown in Fig. 3.

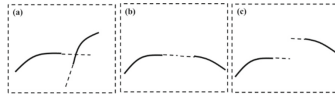


Fig. 2. Track segment pairs with different relative positions. (The dot line represents the extension of the track segment.)

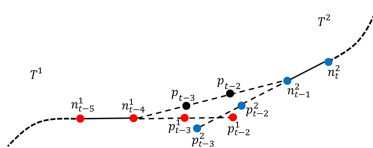


Fig. 3. Similarity measure for trajectory segment merging

T^1 and T^2 do not overlap in time. T^1 ends at frame $(t-4)$ and T^2 starts at frame $(t-1)$. According to the uniform motion constraint, we extend T^1 and T^2 to frame $(t-3)$ and $(t-2)$. The extended target positions are $\{p_{t-3}^1, p_{t-2}^1\}$ and $\{p_{t-3}^2, p_{t-2}^2\}$. If T^1 and T^2 belong to the same target trajectory, we link n_{t-4}^1 and n_{t-1}^2 . Under the uniform motion constraint, the interpolated target positions are $\{p_{t-3}, p_{t-2}\}$ as shown in Fig. 5. This paper uses the distances between extended target positions and the interpolated target positions to define the similarity measure $s(T^1, T^2)$ between T^1 and T^2 as Eq. (6).

$$s(T^1, T^2) = \frac{(t-1) - (t-4) - 1}{\sum_{k=t-3}^{t-2} (\|p_k^1 - p_k\|_2 + \|p_k^2 - p_k\|_2)} \tag{6}$$

The above is a detail description of the similarity definition of two trajectory segments T^1 and T^2 with a time interval 2. For other cases, the similarity is calculated similar to

the above definition. The similarity matrix $S = [s(T^i, T^j)]_{M \times M}$ is a symmetric matrix ($S(i, j) = S(j, i)$) with zero diagonal elements ($S(i, i) = 0$). We get the binary link matrix A_2 as Eq. (7).

$$A_2(i, j) = \begin{cases} 1, & S(i, j) \geq \sigma_3 \\ 0, & S(i, j) < \sigma_3 \end{cases} \quad (7)$$

Where $\sigma_3 = 0.1$ is a set threshold. The parameter corresponds to the degree of relaxation of the long-loose constraint. Considering the continuous of the target movement, we give priority to merging long trajectory segments. We sort the trajectory segments according to their lengths in a descend order, i.e. the first row A_2 in corresponding to the longest trajectory segment in $\{T^i\}_M$.

We delete the trajectory segment without growth or merging in last 5 frames from the list.

After trajectory segment growth and merging, we detect small aerial infrared target according to the length of the trajectory. The length threshold σ_5 is defined as Eq. (8).

$$\sigma_5 = \lfloor \mu \times L \rfloor \quad (8)$$

Where μ is a constant. L is the length of the time window. If the length of a trajectory is larger than σ_5 , the corresponded target candidates are detected as true targets. The target can be detected continuously through the trajectory segment growth. We assume that there is at most one target in a position at the same time. Thus for the crossed trajectory segments, we keep the longest trajectory and eliminate others.

4 Experiments

4.1 Experimental Settings

To validate the performance of the proposed algorithm qualitatively and quantitatively, we conduct experiments on the public datasets SIATD [17]. We perform comparisons between our method with representative existing methods, including single frame based methods (RIPT [18], MLCM [19], AGADM [20]) and multiple frames based method (TIPI [21]). For the compared methods, we use the implementations released by the authors and default parameter settings suggested in their papers. For the proposed method, the parameters are set as follows: local region size $R_1 \times R_2$: 3×3 , 7×7 , 11×11 , 15×15 , length percentage μ : 0.15, and time window length L : 20. For conveniently, we denote our target candidate detection algorithm as ‘‘Med-LGBM’’ and the whole method as ‘‘Proposed’’.

We appoint that the target is correctly detected if the detection locates in the 3-pixel neighborhood of the ground truth in the experiments within this paper.

4.2 Target Detection from Image Sequence

In this section, we conduct experiments of small infrared aerial target detection from image sequences in the SIATD dataset. The proposed method firstly detects target candidates from each single frame via interesting pixels detection and a trained LightGBM

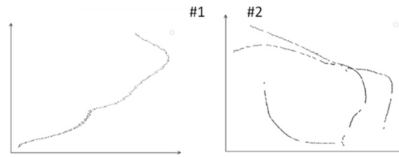
Table 1. Target candidate detection on the testing subset of SIATD dataset.

Algorithm	Recall	Precision	$F_\beta - measure$	Average number of target candidate per frame
Inter-frame differencing	0.73	0.30	0.43	3.96
Med-LGBM	0.74	0.84	0.79	0.87

model as described in Sect. 3.1. For successive frames, a simple commonly used target candidate detection method is inter-frame differencing and thresholding. We compare the simple method with the proposed target candidate detection method on the testing subset of the SIATD dataset. For the inter-frame differencing based method, we perform inter-frame registration as described in Sect. 3.2. And the threshold is determined adaptively via Otsu [22]. The LightGBM model trained on the training subset of the SIATD dataset. The results are presented in Table 1.

The results in Table 1 show that the proposed target candidate detection method achieves higher recall and precision than the inter-frame differencing based method. It is about 0.87 target candidates on average detected by the proposed method which makes the following trajectory growth and merging as efficient as possible. It should be noted that there are at most 3 true targets in each frame and the target may move out of view in SIATD. Thus the average number of target candidate per frame of Med-LGBM is less than 1. The low average number indicates the high precision of the proposed target candidate detection method to some extent.

This paper detects the true targets from target candidates using trajectory constraints. The short-strict and long-loose constraints described in Sect. 3.2 make the proposed method enable to track long target trajectories. We present two detected trajectories within two sample image sequences from SIATD in Fig. 4. It can be seen from Fig. 4 that the targets' trajectories are tortuous but smooth and continuous. The proposed method detects them correctly.

**Fig. 4.** The detected trajectories for two sample image sequences from SIATD dataset.

We train the LightGBM model using the training subset in SIATD. For fairness, we only reported the results on the testing subset in this section. Figure 5 shows sample detection results from the SIATD dataset.

We note that only the correct detections are labeled for the compared methods. Due to too many false alarms detected by some compared methods (e.g. the 2nd-th row in Fig. 5), this paper don't labeled the false detections for clearly. However, there are not much false detection outputted by our method and we labeled them in presented

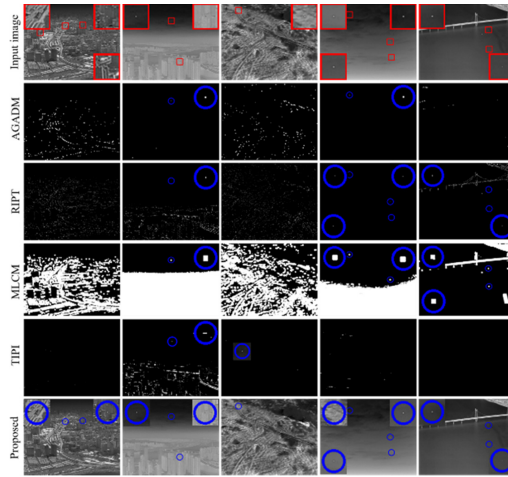


Fig. 5. Sample detection results of AGADM, RIPT, MLCM, TIPI and the proposed method. (The detected target are labeled by “O” and the true targets are labeled by “□”.)

results. The results in Fig. 5 show that existing algorithms have trouble in detecting the small infrared aerial targets, especially for targets under clutter backgrounds. Clutter backgrounds bring great difficulties for AGADM, RIPT and MLCM. False targets are detected in the clutter background area as shown in Fig. 5. TIPI has trouble in modeling

the quick changing of the background. Strong edges bring false detections as shown in the 3rd row in Fig. 5. While the proposed method detects the targets correctly and performs better than other algorithms. As shown in the 2nd column in Fig. 5, a darker target locates in the building region. The proposed method detects it correctly while others not.

For quantitative evaluation, the quantitative evaluations of each algorithm on the testing subset are reported in Table 2, including results of each scene type and the whole testing subset.

Table 2. Quantitative evaluation of AGADM, RIPT, MLCM, TIPI and the proposed method on the testing subset of SATD. (**P** for precision, **R** for recall and **F** for F_{β} – measure.)

		Up looking	Head up looking	Down looking (Vegetation)	Down looking (Water)	Down looking (Building)	Whole
AGADM	P	0.14	0.14	0.03	0.01	0.01	0.01
	R	0.33	0.33	0.18	0.07	0.05	0.02
	F	0.20	0.20	0.06	0.02	0.02	0.01
RIPT	P	0.33	0.33	0.26	0.32	0.30	0.30
	R	0.79	0.79	0.63	0.78	0.74	0.75
	F	0.47	0.47	0.38	0.46	0.44	0.43

(continued)

Table 2. (continued)

		Up looking	Head up looking	Down looking (Vegetation)	Down looking (Water)	Down looking (Building)	Whole
MLCM	P	0.26	0.26	0.14	0.13	0.13	0.05
	R	0.61	0.61	0.33	0.33	0.33	0.13
	F	0.36	0.36	0.20	0.19	0.20	0.08
TIPI	P	0.03	0.01	0.01	0.01	0.00	0.01
	R	0.37	0.02	0.01	0.01	0.00	0.08
	F	0.07	0.01	0.01	0.01	0.00	0.02
Med-LGBM	P	0.54	0.42	0.87	0.50	0.68	0.68
	R	0.87	0.85	0.72	0.88	0.59	0.76
	F	0.67	0.56	0.79	0.64	0.64	0.72
Proposed	P	0.95	0.76	0.99	0.93	0.98	0.92
	R	0.66	0.65	0.60	0.69	0.57	0.63
	F	0.78	0.72	0.75	0.79	0.72	0.75

As mentioned above, clutter backgrounds bring great challenges for existing single frame based methods. AGADM, RIPT and MLCM achieve lower precisions than the proposed method. The cluttered degrees of the backgrounds within the Down looking scene are higher than that within the Up looking scene generally. The performances of the compared methods decrease with the increase of the cluttered degree as shown in Table 2. Our target candidate detection method Med-LGBM achieves better performance than them. LightGBM is a learning based method. Its performance heavily depends on the training data. We reported the results for each scene type separately and the whole testing subset in Table 2. The results on the whole testing subset are similar to those on different scene types. It indicates that Med-LGBM can deal with a variety of complex scenes.

TIPI cannot handle quick changing clutter background well and achieves poor performance on SIATD dataset as shown in Table 2. By introducing the short-strict and long-loose trajectory constraints, the proposed method achieves superior performance than other existing algorithms in detecting small infrared aerial targets as shown in Table 2.

The results in Table 2 indicate that the precision is improved greatly from Med-LGBM to Proposed by introducing trajectory constraints. The false detections are removed effectively. The clutter within the background cannot form a smooth and continuous trajectory as the true target does. However, as Table 2 shows, the recall is decreased from Med-LGBM to Proposed. It means that some correct detected targets are removed wrongly in trajectory growth and merging. We analyzed the experimental results and found that most of the removed correct detected target candidates are isolated detected targets. There is no detected target close to them in the previous and the subsequent

frames. So they cannot form valid trajectory segments and are removed from the final detection.

5 Conclusions

This paper tackles the challenge of small infrared aerial target detection. According to the characteristics of continuity and smoothness of target trajectory, a novel small infrared aerial target detection method using spatial and temporal cues is proposed in this paper. For the target candidate detection from each single frame, using the spatial cues, this paper treats it as a binary classification problem. We use interesting pixel detection and a trained LightGBM model to detect target candidates. For the temporal cues, we adopt the piecewise uniform motion model to approximate the target movement. The true targets are detected from the target candidates using the short-strict and long-loose constraints. The constraints are used in trajectory segment growth and merging. Experiments on the publicly available dataset SIATD indicate that the proposed method achieves better performance than other existing methods.

Acknowledgments. This work is funded by Postgraduate Scientific Research Innovation Project of Hunan Province (project number: CX20200088), Postgraduate Scientific Research Innovation Project of Hunan Province (project number: CX20200024), Postgraduate Scientific Research Innovation Project of Hunan Province (project number: CX20200025) and the National Natural Science Foundation of China (Grant No. 62003357).

References

1. Sun, X., Liu, X., Tang, Z.: Real-time visual enhancement for infrared small dim targets in video. *Infrared. Phys. Tech.* 83, 217–226 (2017)
2. Ke, G., Meng, Q., Finley, T.: LightGBM: a highly efficient gradient boosting decision tree. In: 31st Annual Conference on Neural Information Processing Systems (NIPS), vol. 30. Neural Information Processing Systems (Nips), 10010 North Torrey Pines Rd, La Jolla, California 92037 USA (2017)
3. Moradi, S., Moallem, P., Sabahi, M.F.: A false-alarm aware methodology to develop robust and efficient multi-scale infrared small target detection algorithm. *INFRARED. PHYS. TECHN* 89, 387–397(2018).
4. Gao, C., Meng, D., Yang, Y.: Infrared patch-image model for small target detection in a single image. *IEEE TIP* 22(12), 4996–5009 (2013)
5. Xue, W., Qi, J., Shao, G., Xiao, Z.: Low-rank approximation and multiple sparse constraints modeling for infrared low-flying fixed-wing UAV detection. *IEEE. J-Stars* 99, 1 (2021)
6. Zhao, M., Li, L., Li, W., Li, L.: Infrared small-target detection based on multiple morphological profiles. *IEEE. TGRS* 1–15 (2020)
7. Dai, Y., Wu, Y., Zhou, F., Barnard, K.: Attentional local contrast networks for infrared small target detection. *IEEE. TGRS* 1–12 (2021)
8. Zhao, B., Wang, C., Fu, Q., Han, Z.: A novel pattern for infrared small target detection with generative adversarial network. *IEEE. TGRS* 1–12 (2020)
9. Rodriguez-Blanco, M., Golikov, V.: Multiframe GLRT-based adaptive detection of multipixel targets on a sea surface. *IEEE. J-STARs* 9(12), 5506–5512 (2016)

10. Fan, X., Xu, Z., Zhang, J.: Dim small target detection based on high-order cumulant of motion estimation. *Infrared. Phys. Tech.* **99**, 86–101 (2019)
11. Li, Y., Tan, Y., Li, H., Li, T., Tian, J.: Biologically inspired multilevel approach for multiple moving targets detection from airborne forward-looking infrared sequences. *J. Opt. Soc. Am. A* **31**(4), 734–744 (2014)
12. Zhu, H., Liu, S., Deng, L., Li, Y., Xiao, F.: Infrared small target detection via low-rank tensor completion with top-hat regularization. *IEEE. TGRS* **58**(2), 1004–1016 (2020)
13. Liu, H.K., Zhang, L., Huang, H.: Small target detection in infrared videos based on spatio-temporal tensor model. *IEEE. TGRS* **58**(12), 8689–8700 (2020)
14. Sun, Y., Yang, J., Li, M., An, W.: Infrared small target detection via spatial-temporal infrared patch-tensor model and weighted Schatten p -norm minimization. *Infrared. Phys. Tech.* **102**, 103050 (2019)
15. Sun, Y., Yang, J., An, W.: Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model. *IEEE. TGRS* **99**, 1–16 (2020)
16. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. *Comput. Vis. Image Understand.* **110**(3), 346–359 (2008)
17. SIATD download. <https://small-infrared-aerial-target-detection.grand-challenge.org/Download/>. Accessed 04 Nov 2021
18. Dai, Y., Wu, Y.: Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE. J-STARS* **10**(8), 3752–3767 (2017)
19. Yao, S., Chang, Y., Qin, X.: A coarse-to-fine method for infrared small target detection. *IEEE. GRSL* **16**(2), 256–260 (2019)
20. Wang, G., Zhang, T., Wei, L., Sang, N.: Efficient method for multiscale small target detection from a natural scene. *Optical Eng.* **35**(3), 761–768 (1996)
21. Gao, C., Wang, L., Xiao, Y., Zhao, Q.: Infrared small-dim target detection based on Markov random field guided noise modeling. *Pattern. Recognit* **76**, 463–475 (2018)
22. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE. T. Syst. Man. CY-S* **9**(1), 62–66 (1979)