Infrared small target detection based on region proposal and CNN classifier

Mingming Fan¹ · Shaoqing Tian¹ · Kai Liu¹ · Jiaxin Zhao² · Yunsong Li¹

Received: 11 April 2020 / Revised: 26 March 2021 / Accepted: 13 May 2021 / Published online: 12 June 2021
© The Author(s) 2021

Abstract

Infrared small target detection has been a challenging task due to the weak radiation intensity of targets and the complexity of the background. Traditional methods using hand-designed features are usually effective for specific background and have the problems of low detection rate and high false alarm rate in complex infrared scene. In order to fully exploit the features of infrared image, this paper proposes an infrared small target detection method based on region proposal and convolution neural network. Firstly, the small target intensity is enhanced according to the local intensity characteristics. Then, potential target regions are proposed by corner detection to ensure high detection rate of the method. Finally, the potential target regions are fed into the classifier based on convolutional neural network to eliminate the non-target regions, which can effectively suppress the complex background clutter. Extensive experiments demonstrate that the proposed method can effectively reduce the false alarm rate, and outperform other state-of-the-art methods in terms of subjective visual impression and quantitative evaluation metrics.

Keywords Local intensity enhancement · Corner detection · CNN classifier · Infrared image · Small target detection

1 Introduction

Infrared small target detection is one of the research hotspots in the field of military reconnaissance, which is widely used in the early-warning system, precision-guided weapon, missile tracking system, and maritime surveillance system. Based on the infrared target monitoring system, it uses the infrared radiation differences between background and target to detect the target. Comparing with the radar detection system, the infrared target detection system has strong concealment, easy portability and can detect radar blind area. However, owing to atmospheric disturbance, optical scattering and diffraction, the target has low radiation intensity, low signal-to-clutter ratio (SCR) and lack of shape and texture information. It thus easily drowns into the background [1]. Besides, complex natural scenes such as cloud edges and waves usually introduce false alarm in detection. Therefore,

infrared small target detection is a valuable task with great challenges.

In this paper, the goals of infrared small target detection are to improve the detection rate and to reduce the false alarm rate. In general, infrared small target detection methods are mainly divided into two categories: one is track before detect (TBD) method based on sequence images, and the other is detect before track (DBT) method based on a single-frame image.

TBD methods detect the targets by processing several adjacent frames, which is suitable for the situation that the background changes slowly and the target track is continuous in the infrared sequence. The classical methods are three-dimensional matched filter [2], dynamic-programming method [3], multilevel-hypothesis-testing method [4]. However, multi-frame detection is hard to realize in hardware because of its computational complexity and memory consumption. On the contrary, the DBT method is real-time, low in complexity and easy to implement. It preprocesses the single-frame image, extracts suspicious targets by threshold segmentation, and confirms the targets according to the moving track of the targets on the image sequence. Over the past few decades, many methods based on single-frame detection are proposed.

Filter-based methods use different filter templates to suppress background, such as Max–median/Max–mean

✉ Kai Liu
kailiu@mail.xidian.edu.cn

¹ School of Telecommunications Engineering, Xidian University, Xi'an 710071, China

² Changchun Changguang Insight Vision Optoelectronic Technology Co., Ltd, Changchun 130102, China

filters [5], morphology filtering [6], Top-Hat transform [7], two-dimensional least mean square (TDLMS) filter [8]. It is hard to obtain the universal template parameters and organize elements without prior information, meaning that adaptability and robustness are hard to be guaranteed.

Laplacian of Gaussian (LoG) filter-based method [9] has led to a wide range of research on the human visual system (HVS), Chen [10] proposed a method based on local contrast measure (LCM) for the first time. Then, the improved LCM (ILCM) [11] and the novel local contrast method (NLCM) [12] are proposed to further improve the detection result. Multiscale patch-based contrast measure (MPCM) [13] and multiscale local homogeneity measure (MLHM) [14] detect small targets by enhancing the contrast between the target and its surrounding background. However, methods based on HVS will generate a large number of false alarms when the complex background is complex. Furthermore, Zhang [15] and Liu [16] both utilize the differences between the gradient characteristics of target and background. On the one hand, they can effectively suppress sharp edges. On the other hand, it is time-consuming to calculate the gradient information.

Most of the traditional methods detect targets by hand-designed features, which are usually only effective for a certain background. However, the real infrared small target detection scenes include sky, cloud, sea and various building backgrounds. The traditional method using only artificial features obviously lacks robustness in the complex and changeable application scenes, while the method based on deep learning can extract features by convolutional neural network (CNN), which can make up for the lack of feature and the difficulty of feature description in infrared image.

Deep learning has made great progress in the field of target detection recently, there are mainly two types of detection pipelines: two-stage region-based method [17–19] and one-stage method [20–22]. Wu [23] draws on the experience of the former and applies it to infrared target detection. The potential target region is extracted by the sliding window method, and the real target is judged by a classification network. But it's worth noting that the sliding window method is an actually exhaustive search, which is extremely inefficient. As the lack of public infrared dataset, paper [24] and [25] propose to simulate the infrared image characteristics by visual image data set.

Based on the above observations, a two-stage region-based method is proposed in this paper, using local intensity enhancement and corner detection to find as many real targets as possible, and reducing the false alarm rate by CNN classifier. Our contributions can be summarized as follows:

1. An improved region proposal method is proposed. Instead of sliding window search it improves the efficiency of candidate region search.

2. A lightweight CNN classifier is designed to classify the infrared images with few features and small sizes.
3. A public and diverse infrared image set is established for the study of infrared small target detection.

2 Methodology

Our infrared small target detection method is composed of two modules. The first module takes an infrared image of any size as input. Firstly, enhance the local intensity of the small target; secondly, the coordinates of the suspected target points are found by corner detection; finally, select image patches centered on the coordinates as the potential target regions.

The inputs of the second module are the potential target regions. Each potential target region is fed into the CNN classifier to obtain the category probabilities of potential target region. The category of the region depends on the maximum probability. If it is a target, set the central pixel value of this region to the probability value, and the rest pixel values are set to zeros.

The pipeline of our framework is shown in Fig. 1.

2.1 Region proposal method

2.1.1 Local intensity enhancement

The infrared image generates from the infrared radiation differences between target and background. Infrared small

Fig. 1 Framework overview

targets, such as aircraft and cruise missiles, are thermal radiation sources. Hence, infrared small targets are usually considered to be brighter than surrounding background [15]. In other words, the gray value of small target is larger than its surrounding area, so we perform the local intensity enhancement according to Formula (1). For an infrared small target, the gray value after enhancement will be greater than 0. On the contrary, background pixels with gray values lower than the local mean value will be suppressed. This widens the distinction between small target and background.

$$I(x, y) = \max(0, f(x, y) - \mu), \quad (1)$$

where $I(x, y)$ is the gray value after enhancement, $f(x, y)$ is the gray value at (x, y) , the local mean value μ is the average value of the image patch. Since the size of image patch should be larger than the target size, we adjust the patch size between $7 \times 7 \sim 15 \times 15$ to fit most cases.

2.1.2 Corner detection

A corner is a point around which the pixels vary greatly in gradient amplitude and direction. Qi et al. proposed that the shape of small infrared target is similar to the isotropic Gaussian intensity function due to the optics point spread function (PSF) of the thermal imaging system at a long distance [29]. Figure 2 illustrates five typical examples and their gradient vector graphs. Hence, infrared small target is also a kind of corner.

The eigenvalues of the covariance matrix can be used to distinguish targets from other regions [26]. The eigenvalues depend on the distribution of gradients, so the gradient of the image is analyzed at first. Except small targets, there are also flat regions such as sky background and edge regions such as cloud clutter in an infrared image. The horizontal and vertical gradient distributions of different regions are shown in Fig. 3.

From Fig. 3, the gradients of the pixels in the flat regions are distributed near the origin, with different directions and small amplitudes. The gradients of the edge regions are mainly distributed in a certain direction, which depends on the direction of the edge. However, the gradients of the corners are scattered, not limited to a certain direction and amplitude. Based on the analysis above, if it is a flat region,

Fig. 2 Gradient vector graphs of infrared small targets. The length and direction of the arrow indicate the amplitude and direction of the gradient

then the values of λ_1 and λ_2 are both small, if it is an edge region, then $\lambda_1 \gg \lambda_2$ or $\lambda_2 \gg \lambda_1$, and if it is a corner, both λ_1 and λ_2 are large. Considering the complexity of eigenvalue calculation, the determinant and trace of the covariance matrix are used to measure eigenvalues. The process of region proposal method summarized in Algorithm 1.

Algorithm 1: The process of region proposal method

Input: Infrared image. Parameters: k, t .

The range of k is usually $[0.04, 0.06]$, t is usually around 0.01.

Output: candidate regions

1. **for** $x=1:M$ **do**
 2. **for** $y=1:N$ **do**
 3. Obtain the enhanced gray value $I(x, y)$ according to formula (1)
 4. **end for**
 5. **end for**
 6. **for** $x=2:M-1$ **do**
 7. **for** $y=2:N-1$ **do**
 8. Calculate the gradients in X direction and Y direction respectively,

$$I_x(x, y) = I(x, y) \otimes (-1 \ 0 \ -1),$$

$$I_y(x, y) = I(x, y) \otimes (-1 \ 0 \ -1)^T$$
 9. Calculate the matrix M of image patch W centered on (x, y) , $g(\bullet)$ is Gaussian filtering.

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}, \quad A = \sum_{(x,y) \in W} g(I_x^2),$$

$$B = \sum_{(x,y) \in W} g(I_x I_y), \quad C = \sum_{(x,y) \in W} g(I_y^2)$$
 10. Calculate response value R of each pixel.

$$R = \det(M) - k(\text{trace}(M))^2$$
 11. **end for**
 12. **end for**
 13. Get the set of R larger than the threshold,

$$R = \{R > t \times \max(R)\}$$
 14. Obtain the set of point coordinates corresponding to R ,

$$C = \{(x_i, y_i)\}_{i=1}^K$$
 15. The image patch centered on the obtained coordinates in the input image is taken as the candidate region, the candidate region size is set to 32×32 .
-

2.2 CNN classifier

Small infrared targets can be detected by region proposal method, however, the background may also contain corners.

Fig. 3 Horizontal and vertical gradient distribution of different regions. Each ‘*’ represents the gradient of a pixel, the abscissa represents the horizontal gradient, and the ordinate represents the vertical gradient (A: target region; B, C: edge region; D: flat region)

To eliminate false targets, a lightweight CNN classifier is designed to determine whether the candidate region contains the real target or not.

The proposed classification network model is shown in Fig. 4. The model consists of an image preprocessing layer, a feature extraction module, a global average pooling layer, and a softmax layer.

Even if we collect a large number of infrared images, the overfitting problem still needs to be avoided. Therefore, before the image is fed into the convolution layers, one of the following methods for preprocessing is applied on-the-fly: (a) Horizontal flipping; (b) Contrast adjustment in the range [0.5, 1.5]; (c) Brightness adjustment in the range [−0.12, 0.12]. In this way, the model can see slightly different images in different epoch.

The feature extraction module is composed of convolution layers and pooling layers. A stack of two 3×3 convolution layers is used, which has the same effective receptive field as a 5×5 filter kernel [27]. Feature extraction is performed using two 8-channel convolutional layers at first. After each down-sampling, the number of the feature map channels increased by a factor of 2. The maximum pooling is selected because it is beneficial to retain the gray feature of the infrared small target. The kernel size and stride of the pooling layer are two. Due to the number of categories is 4, a bottleneck layer Conv1 of 4 channels is used to reduce the number of feature map channels to 4. Thus, output of the feature extraction module is a feature map of 8×8 pixels and 4 channels.

Rather than classical fully connected layers after feature extraction, we use global average pooling. There are two beneficial reasons: The fully connected layers have a large number of training parameters and are prone to overfitting. However, in the global average pooling, no parameter needs to be optimized. Furthermore, global average pooling sums out the spatial information that will improve the robustness against disrupted inputs. Global average pooling layer makes

each feature map correspond to a category and outputs the average value of each feature map as a response. Finally, all the responses are fed into the softmax layer for classification. It is identified as a potential target area when the probability of the target category is greater than a certain threshold.

3 Experiment

3.1 Dataset

Infrared images are collected by an infrared camera in summer and winter, respectively, covering a variety of weather and light conditions, as well as various sizes of detected objects. This infrared data set will be open source and will continue to be expanded and enriched.

The size of the collected infrared image is 640 by 480, which does not meet the input requirements of the classifier. In order to obtain the training images, the region proposal method is used to extract potential target regions by cropping patches at 32×32 . Based on our observation, it is inevitable that these regions will contain non-target regions. We reduce the threshold appropriately, in this way, balanced positive and negative samples can be obtained at the same time. Besides, negative samples are divided into three categories because the characteristics of different backgrounds are obviously different. Finally, the total number of dataset images is 28,000, the ratio of target, flat background, complex background and building background is 1:1:1:1, and the training set and validation set are distributed by 7:3. Figure 5 shows some examples of our dataset.

3.2 Proposed CNN classifier training experiment

The CNN classifier is implemented using Python language and TensorFlow package. Training experiments are performed on a computer with an Intel Core i7-7700 CPU, an NVIDIA GPU and 8 GB Memory. We use the following experimental setting for training: the cross-entropy loss is selected as a loss function, and the optimization algorithm used for training is Rmsprop. The other parameters are set in Table 1.

The loss function curve is shown in Fig. 6. It is worth noting that the curve will fluctuate slightly due to the stochastic nature of the learning process. The loss function converges approximately after 8000 iterations. After network training, all the images in the validation set are fed into the model. Finally, the accuracy of validation set is up to 98.11%.

3.3 Comparison experiment and analysis

3.3.1 Evaluation metrics

In order to objectively evaluate the detection performance of different methods. Several widely used evaluation metrics

Fig. 4 CNN classifier for infrared images, a feature extraction module followed by a global average pooling layer, the final category probability is output by softmax

Fig. 5 Some examples of our dataset **a** target, **b** flat background, **c** complex background and **d** building background

including signal-to-clutter ratio (SCR), the background suppression factor (BSF), the receiver operating characteristic (ROC) curve are introduced.

The SCR represents the contrast between target and background, and BSF describes the extent of background suppression, which are defined as follows:

$$\text{SCR} = \frac{f_T - f_B}{\sigma_B} \quad (2)$$

$$\text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}} \quad (3)$$

where f_T is the maximum gray value of the target, f_B and σ_B are mean value and the standard deviation of the local background, respectively, σ_{in} and σ_{out} denote the standard deviation of the original image and the output image, respectively.

Table 1 Training parameter setting

Batch size	128
Weight decay	0.0005
Learning rate	
Initial value	0.001
Updating strategy	Exponential
Decay factor	0.95
Rmsprop optimizer	
Momentum	0.9
Decay	0.9

Fig. 6 Loss function curve (obtained through tensorboard and smoothing coefficient=0.8)

To further evaluate the target detection performance of different methods, the receiver operating characteristic (ROC) curve is also used in this paper. The horizontal axis of the ROC curve is the false alarm rate F_a , the vertical axis is the detection probability P_d . The curve reflects the relationship between the two metrics. The P_d and F_a are defined as follows:

$$P_d = \frac{\text{number of true targets}}{\text{number of all real targets}} \quad (4)$$

Fig. 7 Nine groups of target detection results of the proposed method. The first image of each group is the original image. In the second image of each group, the suspicious target points, in other words, the

center points of the candidate regions, are marked with red circles. The third image of each group shows the three-dimensional result of the proposed method

$$F_a = \frac{\text{number of false targets}}{\text{number of all detected targets}} \quad (5)$$

3.3.2 Comparison with state-of-the-arts

To evaluate the target detection performance of the proposed method, we choose nine public infrared images in various scenes and noisy cases. The target is either submerged in a complex cloud background or exists in a complex sea background. Besides, the background includes not only thick clouds and sheet clouds but also buildings and hills that are prone to cause a high false alarm rate. In the experiments, the parameters in Algorithm 1 are set as $k=0.06$, $t=0.01$ and $n=15$, the threshold of CNN classifier is set to 0.8, and fine-tuning according to the specific situation. As shown in Fig. 7, although the region proposal method recalls the real targets, there are still many false targets in sharp background. However, after the regions are classified, the false alarms can be greatly reduced. Additionally, we find that all the real targets are found with different response values, and only a few false alarms appear in (d) and (e). Experiment results demonstrate that the proposed methods can effectively suppress the background.

To further prove the superiority of the proposed method, six state-of-the-art methods are selected as baseline methods, including multiscale local contrast method (MLCM) [10], multi-patch contrast method (MPCM) [13], average absolute gray difference (AAGD) [28], Laplacian of Gaussian (LoG) filter-based method [9], local intensity and gradient properties (LIG) [15] and flux density and direction diversity in gradient vector field (IMG) [16]. Interference of complex background clutter is always the main difficulty of infrared small target detection, so the ability of background suppression becomes an important standard to judge the performance of a detection method. Figure 8 shows the results of nine images processed by the above six methods and the proposed methods. We can compare the advantages and disadvantages of the above methods intuitively. It can be seen that the MLCM method significantly enlarges the size of each target, but it only enhances the original image roughly and does not suppress the background obviously. MPCM and LIG methods suppress the background to a certain extent, as shown in (b–d) and (h), but they are still sensitive to strong cloud edges in (f–g). AAGD and LoG methods have the same problems, especially, they are vulnerable to the interference of clutters around the sea-sky line and the boundary of hills. IMG method proposes the multiscale flux density that the small targets can be effectively enhanced, background clutters can

Fig. 8 Results and the 3D grayscale images of different methods

Fig. 9 SCR and the BSF of different methods

Fig. 10 ROC curves of three real infrared sequences. **a** The results of all baseline methods and the proposed method. To show the probability of detection with a smaller false alarm rate, the details of ROC curves are shown in **(b)** and **(c)**

be simultaneously suppressed, and the heavy clutter can be further suppressed by the gradient direction diversity, it is indeed superior to other baseline methods in background suppression. Nevertheless, as with other baseline methods, there are still a large number of false alarms left at the building signboard in (i). The proposed method gives satisfactory results, no matter the background is thick cloud, line clutter or building signboard. Accordingly, our method has robust performances for the various scenes.

The above experiments qualitatively demonstrate the excellent performance of the proposed method, the following will compare with baseline methods quantitatively using evaluation metrics. Results of SCR and BSF are shown in Fig. 9. In general, the higher the value of these two metrics means the better the detection performance of the method.

The SCR and BSF of MLCM are not ideal due to the simple calculation, but those of the IMG method are relatively higher. Interestingly, the metric values of MPCM vary greatly under different conditions, which shows its robustness is not good. As can be seen in Fig. 9, the proposed method achieves the highest evaluation metrics in different scenarios, meaning that the proposed method outperforms other state-of-the-art methods.

In order to further evaluate the performance of the proposed method, we tested ROC curves of different sequences. The probability of detection gradually raises with the increase of false alarm rate. In the terms of the ROC curves in Fig. 10, the probabilities of detection of baseline methods can almost reach nearly 100% when the false alarm rate reaches a certain value. However, at the same probability of detection, the

proposed method has the lowest false alarm rate, indicating the proposed method has the best performance in background suppression and target enhancement simultaneously.

4 Conclusion

In this paper, an infrared small target detection method based on candidate region proposal and CNN classifier is proposed, which can separate the real target from the background and greatly reduce the false alarm rate caused by complex background clutter. In particular, after off-line training, the CNN classifier can eliminate special backgrounds such as buildings. Extensive experiments demonstrate that the proposed method has better detection performance than state-of-the-art methods in different scenarios. In future work, we plan to concentrate on the time consumption for computing.

Funding This work was supported by the National Natural Science Foundation of China (Grant No. 61850410523).

Declaration

Conflict of interest All authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Wang, C., Qin, S.: Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis. *Infrared Phys. Technol.* **69**, 123–135 (2015)
- Reed, I.S., Gagliardi, R.M., Stotts, L.B.: Optical moving target detection with 3-D matched filtering. *IEEE Trans. Aero Electron. Syst.* **24**(4), 327–336 (1988)
- Grossi, E., Lops, M., Venturino, L.: A novel dynamic programming algorithm for track-before-detect in radar systems. *IEEE Trans. Signal Process.* **61**(10), 2608–2619 (2013)
- Liu, B., Lan, J., Li, X.R.: A 2-SPRT based approach to multiple-model hypothesis testing for multi-distribution detection. *IEEE Trans. Signal Process.* **64**(12), 3221–3236 (2016)
- Deshpande, S.D., Er, M.H., Venkateswarlu, R., Chan, P.: Max-mean and max-median filters for detection of small targets. *Proc. SPIE* **3809**, 74–83 (1999)
- Ye, B., Peng, J.: Application of order morphology filtering on detection of small target and point target. *Proc. SPIE* **4554**, 94–99 (2001)
- Deng, L., Zhu, H., Zhou, Q., Li, Y.S.: Adaptive top-hat filter based on quantum genetic algorithm for infrared small target detection. *Multimed. Tools Appl.* **77**(9), 10539–10551 (2018)
- Cao, Y., Liu, R.M., Yang, J.: Small target detection using two-dimensional least mean square (TDLMS) filter based on neighborhood analysis. *J. Infrared Millimeter Waves* **29**(2), 188–200 (2008)
- Kim, S., Yang, Y., Lee, J.: Small target detection utilizing robust methods of the human visual system for IRST. *J. Infrared Millimeter Waves* **30**(9), 994–1011 (2009)
- Chen, C.L.P., Li, H., Wei, T.Y.: A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **52**(1), 574–581 (2014)
- Han, J., Ma, Y., Zhou, B.: A robust infrared small target detection algorithm based on human visual system. *IEEE Geosci. Remote Sens.* **11**(12), 2168–2172 (2014)
- Qin, Y., Li, B.: Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci. Remote Sens.* **13**(12), 1890–1894 (2016)
- Wei, Y., You, X., Li, H.: Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recogn.* **58**, 216–226 (2016)
- Nie, J., Qu, S., Wei, Y.: An infrared small target detection method based on multiscale local homogeneity measure. *Infrared Phys. Technol.* **90**, 186–194 (2018)
- Zhang, H., Zhang, L., Yuan, D.: Infrared small target detection based on local intensity and gradient properties. *Infrared Phys. Technol.* **89**, 88–96 (2018)
- Liu, D., Cao, L., Li, Z.: Infrared small target detection based on flux density and direction diversity in gradient vector field. *IEEE J. Stars* **11**(7), 2528–2554 (2018)
- Girshick, R., Donahue, J., Darrell, T.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2014), pp. 580–587.
- Girshick, R.: Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2015), pp. 1440–1448.
- Ren, S., He, K., Girshick, R.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: *Proceedings of Advances in Neural Information Processing Systems* (Academic, 2015), pp. 91–99.
- Redmon, J., Divvala, S., Girshick, R.: You only look once: unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 779–788.
- Liu, W., Anguelov, D., Erhan, D.: SSD: Single shot multibox detector. In: *Proceedings of European Conference on Computer Vision* (Academic, 2016), pp. 21–37.
- Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 7263–7271.
- Shuang-Chen, W., Zheng-Rong, Z.: Small target detection in infrared images using deep convolutional neural networks. *J. Infrared Millimeter Waves* **38**(3), 03371 (2019)
- Fan, Z., Bi, D., Xiong, L.: Dim infrared image enhancement based on convolutional neural network. *Neurocomputing* **272**, 396–404 (2018)
- Wang, W.T., Qin, H., Cheng, W.: Small target detection in infrared image using convolutional neural networks. *Proc. SPIE* **10462**, 1046250 (2017)
- Harris, C.G., Stephens, M.: A combined corner and edge detector. Academic Press, London, pp. 10–5244 (1988)

27. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:[1409.1556](https://arxiv.org/abs/1409.1556) (2014)
28. Deng, H., Sun, X., Liu, M.: Infrared small-target detection using multiscale gray difference weighted image entropy. *IEEE Trans. Aero. Electr. Syst.* **52**(1), 60–72 (2016)
29. Qi, S., Ma, J., Tao, C., et al.: A robust directional saliency-based method for infrared small-target detection under various complex backgrounds. *IEEE Geosci. Remote Sens. Lett.* **10**(3), 495–499 (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.