Generative Adversarial Differential Analysis for Infrared Small Target Detection

Zongfang Ma, Shuo Pang, and Fan Hao

***Abstract*—Infrared small target detection refers to the extraction of small targets in a complex, low signal-to-noise ratio background. Depthwise convolution makes it difficult to comprehensively characterize small infrared targets and ignores the importance of image background for the detection task. In this article, we propose the generative adversarial differential analysis (GADA) model for infrared small target detection, the core of which aims to weaken the reliance on target features and enhance the use of background information. Specifically, we first construct pseudobackground labels by the fast marching method. Then, the background-guided generative adversarial network is used to learn the background data distribution. On this basis, the differential image containing interest regions of small targets is obtained by differential analysis. Finally, the detection results are obtained by performing an elaborate characterization of the interest regions. The effectiveness of GADA is verified with three public datasets. Compared to several state-of-the-art methods, GADA achieves better performance in terms of $F1$, IoU, and AUC.**

***Index Terms*—Change detection, generative adversarial network (GAN), image segmentation, infrared small target detection.**

I. Introduction

A. Background

SINGLE-FRAME infrared small target detection technology is used in several applications, including airborne early warning, maritime surveillance, precision guidance, and other fields [1], [2], [3]. Infrared images are imaged by the infrared system to detect targets in a scene at a long distance. Therefore, compared with general target detection tasks, infrared small targets have the following characteristics: 1) Infrared small targets are generally very small, with pixels occupying only a few or dozens of pixel sizes. For example, a small target is less than 80 pixels in the total spatial extent of a 256×256 image [4]. It results in the lack of information about the shape and texture of infrared small targets; 2) Infrared small targets are usually located in complex backgrounds or receive clutter interference. Therefore, efficient and accurate detection of infrared small targets is still a challenging problem [5], [6]. In recent years, researchers have developed several pioneering works based on image processing and deep learning techniques. These methods can be divided into the following two categories: one is based on model-driven, and the other is based on data-driven.

The model-driven method involves manually designing algorithms in response to assumptions about the physical properties of infrared targets [7]. These methods can be subdivided into three categories: filtering-based methods, human visual system-based methods, and low-rank sparse recovery-based methods. Filter-based methods include Max-median and Max-mean [8], Top-Hat [9], and multiple morphological profiles (MMP) [10], etc. This type of method uses filters to estimate the background of the infrared image to achieve the purpose of suppressing the background or using the frequency difference between the target and the background to filter. The method is better for uniform background suppression but is not as effective for complex backgrounds. Human visual system, such as local contrast measure (LCM) [11] and multiscale patch-based contrast measure (MPCM) [12], halo structure prior-based LCM (HSPLCM) [13], and the idea combining an improved density peak global search and local contrast calculation [14], mainly exploits the intrinsic properties of infrared small targets. The local texture of the image changes when the small target appears instead of the global texture, so the local features are used to complete the detection task. However, it has poor suppression ability for clutter and it is not suitable for low contrast targets. The methods based on low-rank sparse recovery include the infrared patch-image (IPI) model [15] and low-rank and sparse representation (LRSR) model [16], and interpoint correlation enhancement (IPCE) [17] etc. Since infrared small targets have sparse features and the background has low-rank features, the detection task is completed by optimizing the constraints. However, it still has a high false alarm (FA) rate for small targets with different shapes in complex backgrounds [18].

Data-driven models are used to extract the features of various types of data by learning from the data, and further implementing tasks such as classification or detection of them. MDvsFA cGAN [19] decomposes the infrared small object segmentation problem into two subtasks of suppressing MD and FA, respectively, and solving two tasks via generating adversarial learning models. Asymmetric contextual modulation (ACM) [20] proposes an ACM module to combine high-level semantics with low-level features. Dense nested attention network (DNANet) [21] designs a densely nested interaction module for the combination of high-level and low-level features

Manuscript received 13 December 2023; revised 4 February 2024 and 1 March 2024; accepted 3 March 2024. Date of publication 6 March 2024; date of current version 19 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62276207, and in part by the Technology Innovation Leading Program of Shaanxi under Grant 2023GXLH-055. *(Corresponding author: Fan Hao.)*

Zongfang Ma and Shuo Pang are with the School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China (e-mail: zongfangma@xauat.edu.cn; pangshuo@xauat.edu.cn).

Fan Hao is with the School of Integrated Circuits, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: haofan@bupt.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3374054

for progressive interaction and proposes cascade channel and spatial attention modules to enhance multilevel features. Interior attention-aware network (IAANet) [22] obtains the coarse target region and filters the background through a region proposal network, then outputs the attention-aware features before obtaining the prediction results through classification.

B. Motivation

Due to the small size of the target, the data-driven-based methods tend to lose the target feature information when performing deep convolution and also ignore the utilization of the background information. The methods based on principal component analysis point out that an infrared image can be represented as a superposition of background, target, and clutter. Therefore, we would like to construct a network that generates infrared images without target components and then quickly locate small targets by the difference between the generated image and the original image. The generated image and the original image can be viewed as representations of the same region at different times, and the small target is the difference between these two representations. As a result, the infrared small target detection problem can be transformed into a time series change detection problem, avoiding the difficulty of directly extracting small target features from the convolution.

C. Contributions

In this article, we propose a generative adversarial differential analysis network to improve the detection accuracy of infrared small targets. The background-guided generative adversarial network is used to generate the background image of the infrared image. The region of interest containing small targets is obtained by analyzing the difference between the infrared image and the background image through the change detection method. Finally, the detection task is accomplished by precision extraction of the interest region. The main contributions of this work are summarized as follows:

1) To improve the detection performance of infrared small targets, GADA converts the infrared small target detection problem into a time-series change detection problem, which has the advantage of weakening the dependence on target features and enhancing the utilization of background information.
2) The GAN-based generative adversarial differential analysis module and U-Net-based precision extraction module are designed to separate the target from the background image more efficiently and obtain better detection results.
3) To address the problem of missing background images in the infrared image dataset, an image inpainting technique based on the fast marching method is used to obtain a large number of weak labels, which are used to constrain the background distribution of the generated infrared image.

The rest of the article is organized as follows. Section II reviews several related works. Section III describes the details of GADA. Section IV validates the reliability of GADA through comparative experiments and discussion. Finally, Section V concludes the article

II. Related Work

A. Convolutional Generative Adversarial Networks

Generative adversarial network (GAN) is a combination of neural networks and game theory for learning approximate distributions of real data from a large number of sampled inputs. It generally consists of a generator and a discriminator. The generator is used to learn the distribution of real data to generate an image that approximates the distribution of real data, and the discriminator is used to determine whether the image is real data or generated data. In practice, the parameters of the generator and the discriminator alternately change each other, so that they reach a Nash equilibrium, and then no longer change their weights. DCGAN [23] is an extension model of GAN, which is a model proposed by combining CNN and GAN. By adjusting the structure of the network to make the network more stable and have better generalization.

In the infrared small target detection task, Wang et al. [19] designed the MDvsFA cGAN. Using the network architecture of conditional GAN, two generators were designed to minimize miss detection (MD) and FA, respectively. Zhao et al. [24] designed the IRSTD-GAN. This network treats small targets as a special kind of noise, which is predicted from the original image based on the data distribution learned by GAN. In summary, the GAN-based techniques have been demonstrated to be effective for the task of infrared small target detection.

B. Image Inpainting

Image inpainting aims to restore pixel features to damaged parts of a mutilated image. Image inpainting serves a wide range of applications, such as removing text and logos from still images or videos, reconstructing scans of damaged images by removing scratches or stains, or creating artistic effects [25].

The image inpainting technique based on the fast marching method is to process the pixel points at the edges of the area to be repaired first, and then layers are pushed inward until all pixel points have been repaired. The gray values of the pixel point to be repaired need to be calculated from all the points in the domain. This process can be expressed as

$$I(p) = \frac{\sum_{q \in B_\varepsilon(p)} \omega(p,q)[I(q) + \nabla I(q)(p-q)]}{\sum_{q \in B_\varepsilon(p)} \omega(p,q)} \quad (1)$$

where p denotes the pixel point to be repaired, B_ε denotes a small field selected with p as the center, q denotes a known pixel point in the field, and the pixel value of point p is calculated from point q. Different pixel points have different influences on the pixel point to be repaired due to factors such as distance. ω denotes the weight function by which the influence of different pixel points in the domain is limited.

C. Change Detection

Change detection refers to identifying the differences between remote sensing images acquired over the same geographical zone but taken at two distinct times [26]. Classical change detection methods can be constructed based on different operators and algorithms, such as univariate image differencing [27], change

vector analysis (CVA) [28], ratioing [29], distance or similarity measures [30], etc. CVA is able to explore the dissimilarity between features corresponding to changed and unchanged regions.

In GADA, the infrared small target image and the background image are considered as the representation of the same region at different times. Based on this, the infrared small target detection problem is converted into a time series change detection problem. The recognition of change region is a replaceable part of GADA. In recent years, the rapid development of deep learning-based change detection methods has made its application in small target detection frameworks more possible. For example, Samadi et al. [31] proposed a method that combines morphological images with two original images to provide a suitable data source for DBN training. Shu et al. [32] proposed a two-stage patch-based deep learning method using a label update strategy. Initial labels and masks are generated in the preclassification stage. Then a two-stage update strategy is used to recover the changed regions gradually. Du et al. [33] proposed a new end-to-end TransUNet++SAR network, which incorporates a vision module in the network architecture. These methods have led to enlightening explorations in infrared small target detection.

D. Image Segmentation

Image segmentation techniques refer to the division of an image region into two or more meaningful regions or the definition of boundaries for different semantic entities in an image [34]. Image segmentation is a classification task at the pixel level. It essentially consists of an encoder and a decoder. The encoder is used to extract the high-dimensional features of the image. The decoder, on the other hand, generates semantic segmentation masks from the high-dimensional feature vectors by upsampling such as deconvolution, interpolation, etc., and maps them back to the original image.

Infrared small target detection can be considered as a binary semantic segmentation with an extreme imbalance of positive and negative samples [35]. Currently, there are three main methods based on CNN, based on CNN+model-driven and based on CNN+Transformer. Among them, CNN-based methods mainly utilize classical segmentation networks. For example, inspired by the feature extraction of DFN [36] and SENet [37], APGC-Net [7], and AFFPN [38] adopt top-down channel attention mechanism and bottom-up spatial attention mechanism to extract high-level semantic information and low-level feature information respectively. CNN+model-driven methods mainly utilize the difference in local contrast, and the combination of the two improves local contrast or performs local contrast learning. For example, a new multiscale local contrast learning network [39] introduces local contrast learning in the network to learn the local contrast features of small infrared targets. The CNN+transformer-based methods use the self-attention mechanism of the transformer to learn the interaction information of image features in a wider range [35]. IAANet [22] acquires the target coarse region and filters the background, after which the attention between the pixels in the coarse region is modeled by the transformer encoder. In conclusion, the methods based on image segmentation are effective for infrared small target detection.

III. Methodology

A. Overall Architecture

Fig. 1 shows the detailed structure of the GADA model. By feeding a single infrared image into the generative adversarial differential analysis module, we can get the infrared background image and the differential image. Then, the target detection result is obtained by the precision extraction.

Section III-B describes the architecture of the generative adversarial differential analysis module. Feature extraction and upsampling of infrared images by encoding and decoding reconstruct the data distribution of the original image to obtain a background image. The differential image is obtained by the differential operation between the original image and the background image. Section III-C presents a detailed elaboration of the precision extraction module. The difference image is fed into this module to achieve small target detection.

B. Generative Adversarial Differential Analysis

As shown in (2), principal component analysis defines that an infrared image consists of a superposition of a background, a target, and clutter. We would like to construct a network that generates an infrared image without small targets, and then quickly locate the small targets by the difference between the generated image and the original image

$$f_D(x,y) = f_T(x,y) + f_B(x,y) + f_N(x,y) \tag{2}$$

where f_D, f_T, f_B, and f_N denote the input image, the target component image, the background component image, and the noise component image, respectively.

As shown in part 2 of Fig. 1, the generative adversarial network is used to build the generative module. We chose a deep convolutional generative adversarial network, which consists of a U-Net-based generator and a full convolutional discriminator. Using the encoder and decoder of the generator, the multilevel information of the image background is extracted and the data distribution of the background is reconstructed. The reconstructed details of the image are supplemented by the skip connection. The discriminator utilizes real data samples to determine the true probability of the reconstructed image and further adjusts the network parameters.

The real samples required by the generative model are background images without small targets, which are not provided in existing infrared datasets. Therefore, a pseudolabeled dataset is constructed using the fast marching-based image restoration technique mentioned in Section II. The region to be repaired is identified from the labels in the infrared dataset, and the complete background image is obtained using the restoration technique. The restoration operation is shown below:

$$\theta_i = I(p_i\,|_{p_i\in\Omega, i=1,2,\cdots n}) \tag{3}$$

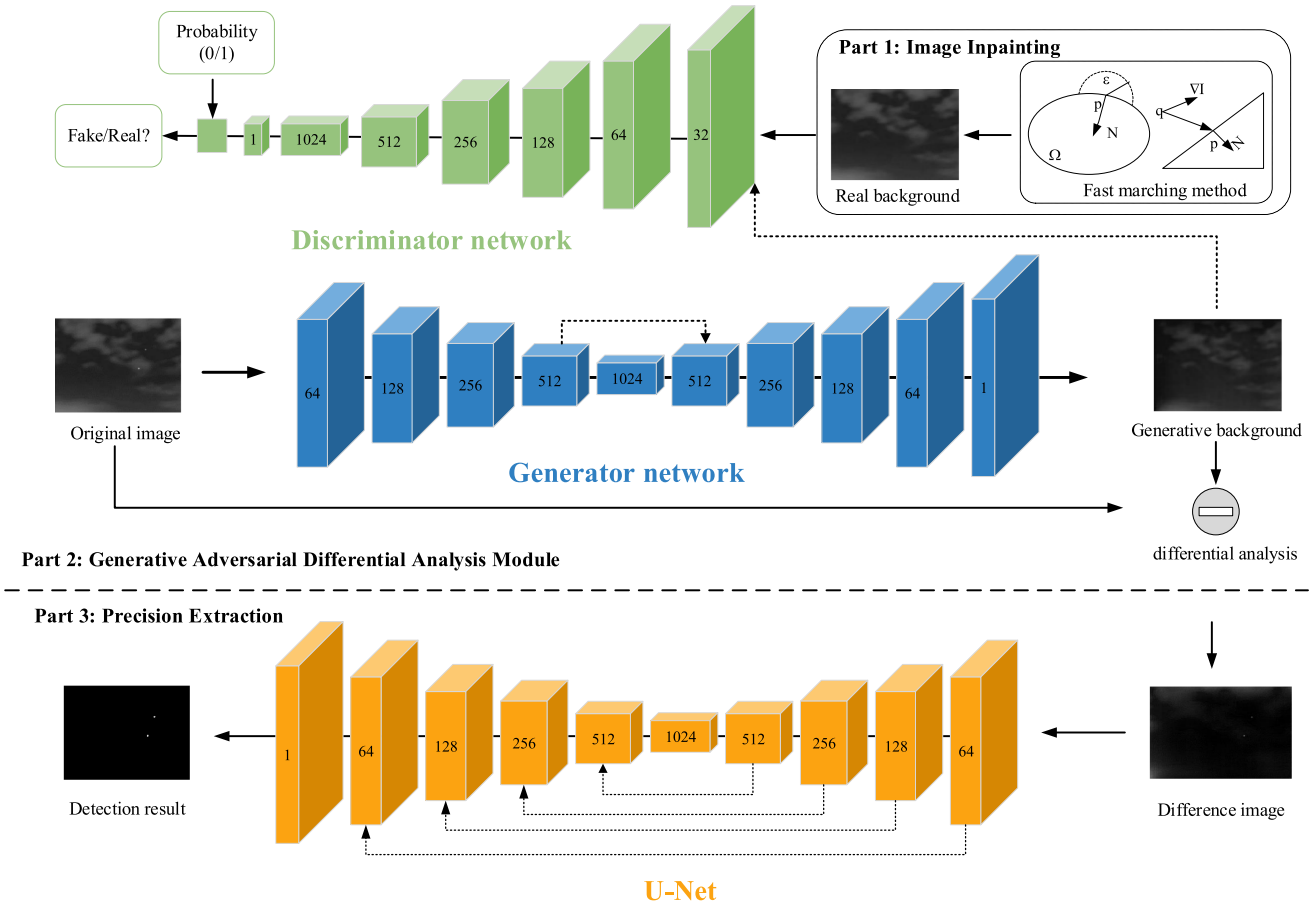

Fig. 1. Overall architecture of the GADA, which contains an image inpainting module, a generative adversarial differential analysis module, and a precision extraction module.

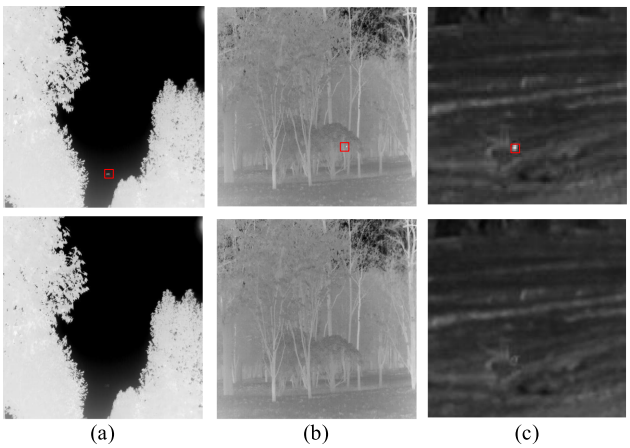

Fig. 2. Original image and the background image obtained by image inpainting.

where Ω denotes the area to be repaired and p_i denotes the ith pixel of the Ω. I denotes the fast marching method. The generated pseudolabel images are shown in Fig. 2.

The generator uses U-Net as the basic model. As a classical segmentation network, U-Net adopts a U-shaped structure and the skip connection to better retain the image information. The input infrared image x_1 is processed through the encoder, which is mainly implemented using convolutional operations. The image is downsampled with stride of 2. The output x_2 is defined as

$$x_2 = \text{Conv}(\delta(B(\text{Conv}(\delta(B(\text{Conv}(x_1)))))))) \tag{4}$$

where Conv denotes the convolution operation with a convolution kernel of 3×3. B denotes the batch normalization. δ denotes the activation function. The LeakyRelu activation function with the negative value set to 0.2 is used by default, but the Tanh activation function is used in the last layer. x_2 to x_6 are the outputs of each layer of the encoder.

The feature map is fed to the decoder and the image size is increased using deconvolution with the stride of 2. The output is defined as

$$x_8 = \text{ConvTrans}(\delta(B(\text{Conv}(\delta(B(\text{Conv}(x_7)))))))), \tag{5}$$

where ConvTrans denotes the ConvTranspose operation. x_7 to x_{12} are the outputs of each layer of the decoder.

The discriminator consists of a fully convolutional network. There are seven layers of convolution and each layer consists of convolution, normalization, and activation function, where the stride is 2 and the last layer outputs the probability of true or false using Sigmoid function.

The generative adversarial differential analysis model contains three parts of loss, which are generative adversarial loss, reconstruction loss, and perceptual loss.

Generative Adversarial Loss: The generative adversarial loss can be expressed as follows:

$$\mathcal{L}_{Gan} = E_{x\sim p_x}[\log D(x)] + E_{z\sim p_z}[\log(1 - D(G(z)))] \quad (6)$$

where p_x represents the distribution of real samples, p_z represents the data distribution of the input image, then G is the generator, and D is the discriminator.

Reconstruction Loss: Reconstruction loss aims to make the generated image close to the real image at the pixel level. It can be expressed as follows:

$$\mathcal{L}_{re} = \min |x - G(z)| \quad (7)$$

where x denotes the real image and $G(z)$ denotes the sample generated by the generator.

Perceptual Loss: In order to make the generated background image as close as possible to the original background image and to focus more on the learning of texture features, perceptual loss [40] is used to make the two more semantically similar. We chose the block3-conv3 layer of the pretrained 19-layer VGG network and extracted the corresponding feature maps for computation. The details are shown as follows:

$$\mathcal{L}_{pre} = \frac{1}{C_j H_j W_j} ||\phi_j(x) - \phi_j(G(z))||_2^2 \quad (8)$$

where j represents the jth layer of the network, CHW represents the size of the extracted feature maps, ϕ represents the selected pretrained network, x represents the real samples, and $G(z)$ represents the generated samples.

So, the total loss function of the generative model can be expressed as

$$\mathcal{L} = \mathcal{L}_{Gan} + \alpha\mathcal{L}_{re} + \beta\mathcal{L}_{pre} \quad (9)$$

where α and β are empirical parameters.

It is observed that there is a large difference between the generated image and the original image in the small target region. Therefore, the pixel-by-pixel differencing method is utilized to obtain the difference between the two images and get the difference image containing the target. The representation is as follows:

$$d = |x_1 - x_{12}| \quad (10)$$

where x_1 denotes the original image and x_{12} denotes the generative image.

C. Precision Extraction of Small Target Regions

Due to the limitations of the generative model, the generated background image is slightly different from the pseudolabel. The difference map contains obvious small targets and a little interference information. It needs to be extracted by fine extraction to get the final detection result. We choose the classical network U-Net.

The encode is realized by convolution operation, which decreases the feature map size sequentially. The dropout operation is added to prevent the overfitting phenomenon. The operation of each layer can be expressed as follows:

$$x_2 = \delta(D(B(\text{Conv}(x_1)))) \quad (11)$$

where D denotes the dropout operation.

The decoder recovers the image details and uses neighborhood interpolation to gradually recover the image size. Supplement the feature information of small targets by the skip connection. The operations can be expressed as follows:

$$z = \delta(B(\text{Conv}(C(U(x), y)))) \quad (12)$$

where x denotes the feature map, y denotes the downsampled feature map corresponding to x, and U denotes the interpolation operation. C denotes skip connection.

Instead of using the L_2 loss function for the segmentation results with the label image, the segmentation model chooses dice loss [41]. Dice loss is a region-dependent loss function used to solve the problem of positive and negative sample imbalance. The larger the coefficient is, the more similar the samples are to each other. The loss function is represented as follows:

$$\mathcal{L}_{\text{dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (13)$$

where $|X|$ and $|Y|$ denote the number of elements in the segmentation result and the label image, respectively, $|X \cap Y|$ denotes the intersection between the segmentation result and the label image.

IV. Experiments

In this section, we evaluate the effectiveness of GADA with $F1$, IoU, and AUC metrics on three datasets. The experimental setup, including the dataset, the baseline methodology, and the evaluation metrics are described. Then, a comparison with the baseline method is made to illustrate the superiority of the GADA model. Finally, the comprehensive performance of the method is verified by some discussion experiments.

A. Experimental Settings

1) Datasets: We conducted experiments on IRSTD-1k [18], SIRST dataset [20], and MSISTD [42], respectively. SIRST contains 427 infrared images. About 55% of the targets in this dataset occupy only 0.02% of the image area, and about 35% of the targets are the brightest [20]. IRSTD-1 k contains 1000 infrared images. These images were captured by infrared cameras in the real world and covered a variety of scenes with backgrounds such as oceans, rivers, fields, mountains, cities, and clouds, with heavy clutter and noise [18]. MSISTD contains 1077 images. It greatly expands the context of single-frame small target detection from being limited to the detection of aerial targets only to the detection of multiscene, multiscale, and lower SCR targets [42]. For three datasets, they are divided into training and testing sets in the ratio of 8:2.

2) Baseline Methods: We compared GADA with focused infrared small target detection methods including new Tophat [9], IPI [15], nonconvex rank approximation minimization (NRAM) [43], ACM [20], improved multimode nuclear norm joint local weighted entropy contrast (IMNN-LWEC) [44], sparse regularization-based spatial–temporal twist tensor

(SRSTT) [45], interior attention-aware network (IAANet) [22], attention-guided pyramid context networks (AGPCNet) [7], and low-level network (ILNet) [46]. Tophat [9] was achieved by filtering to achieve background suppression. IPI [15] formulated the infrared small target detection problem as an optimization problem that recovers a low-rank sparse matrix. NRAM [43] employed a nonconvex, tighter rank surrogate and a weighted L_1 norm to efficiently retain the target while suppressing the background. ACM [20] was based on a deep learning network that combines the high-level semantic information of the small target with the low-level features. IMNN-LWEC [44] represented the infrared target detection task as an optimization problem of tensor decomposition of the three components of the background tensor, the target tensor, and the sparse structure tensor in the spatio-temporal domain. SRSTT [45] increased the difference between the background and the target by twisting the tensor model and combining the structured sparse-induced paradigm and the L_1-paradigm as a parsing constraint for the target. IAANet [22] utilized a region proposal network (PRN) to obtain the coarse target region and filter the background. AGPCNet [7] utilized a designed context module CPM to better adapt to the characteristics of small infrared targets, resulting in a performance boost. ILNet [46] took the important underlying features from shallow to deep layers and dynamically assigned weights to shallow and deep layers.

3) Evaluation Metrics: Infrared small target detection mainly uses pixel-level evaluation metrics such as IoU, $Precision$, and $Recall$. IoU is used to evaluate the accuracy of localization, which is calculated by the ratio of the intersection area to the concatenation area between the prediction and the label. Simply achieving a high Precision or a high Recall cannot indicate whether a method is good or bad, so we select the F-measure [47] metric as the reconciled mean of the two. F1 is calculated as shown in (14). The receiver operation characteristics (ROC) curve reflects the dynamic relationship between the false positive rate and the true positive rate. AUC denotes the area under the curve, which is used to quantitatively evaluate the classification performance of true and false targets [48], [49].

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision}. \tag{14}$$

4) Implementation Details: In the generative adversarial differential analysis module, Adam is selected as the optimizer. The generator learning rate is 0.0001 and the discriminator learning rate is 0.001. Epoch and batch size are set to 30 and 16, respectively. In the precision extraction module, the Adam optimizer is used with the learning rate set to 0.0001, the epoch of training is 40 and the batch size is 16.

B. Comparison to State-of-the-Art Methods

To demonstrate the superiority of GADA, we compare the GADA with several state-of-the-art methods using visual and numerical evaluation results.

1) Visual Evaluation: We selected the detection results of six images from the SIRST, IRSTD-1 k, and MSISTD datasets, respectively, and compared them with nine state-of-the-art methods. The detection results are shown in Figs. 3–5. The red box indicates the correct detection result, the green dashed circle indicates the missed detection region, and the yellow dashed circle indicates the FA region.

Fig. 3 shows the detection results of the dataset SIRST. Image-1 shows a car driving on a highway surrounded by bushes, where the bushes are more similar to the small target features. New Tophat, IPI, and SRSTT are prone to generate FAs in the bushes part. Also, New Tophat, IPI, NRAM, IMNN-LWEC, and SRSTT do not retain the target contour well, and the detection results are incomplete. Image-2 shows the sea surface under sunlight, New Tophat, IPI, and NRAM are prone to generate FAs in the area where the sea surface is illuminated to produce highlights, and IMNN-LWEC has missed detection. Image-3 shows the sky with clouds. New Tophat, IPI, and SRSTT are prone to generate FAs in the highlighted part of the cloud image. Image-4 is a highway, the distant mountain peaks and the trees along the roadside belong to the highlighted area. New Tophat, IPI, NRAM, IMNN-LWEC, SRSTT, and IAANet are prone to generate FAs in the tree part. NRAM and ILNet have missed the detection phenomenon. Image-5 shows the sea surface with waves rising, and New Tophat and IPI are prone to generate FAs in the highlighted area of the waves. AGPCNet has missed the detection phenomenon. Image-6 shows the sky with the pure background. IPI has a small portion of FAs, and IMNN-LWEC has missed detections for targets with weak intensity. The presentation of the results also shows that the traditional method has more complete detection results for point-shaped targets, and there is a problem of partial missing detection results for other shaped targets, and the target contour is not well preserved.

Fig. 4 shows the detection results of the dataset IRSTD-1 k. Image-1 shows a forest with obvious small targets and less interference information. The targets are detected by all methods, but the ACM has FAs near the target point. Image-2 shows a lawn in front of a house with highlighted regions, New Tophat, IPI, and IAANet detect small targets with FAs, and NRAM and SRSTT have missed detections. Image-3 shows a lake with reflections of tall buildings and there are highlighted regions on the lake in the image. New Tophat, IPI, ACM, and IMNN-LWEC have FAs and IMNN-LWEC detects missing targets. Image-4 is a tall building in the city. New Tophat, IPI, NRAM, IMNN-LWEC, ACM, SRSTT, IAANet, and ACPCNet all falsely detect lights in the tall building as targets. AGPCNet have missed the detection phenomenon. The result obtained by ILNet is not obvious. Image-5 shows the open space in front of the bushes, where IPI, NRAM, IMNN-LWEC, SRSTT, IAANet, and AGPCNet have missed detections due to the small size of the target and low contrast with the surroundings, New Tophat and ACM detected the target, but there is a FA in the bushes area. At the same time, the result obtained by ILNet is not clear. Image-6 includes streetlights and a sky containing clouds. The small target is more obvious, but the streetlight is the highlighted area. Therefore New Tophat, IPI, NRAM, IMNN-LWEC, SRSTT, and IAANet all produce FAs in the highlighted region.

Fig. 5 shows the detection results of the dataset MSISTD. Image-1 shows the image captured by looking up at the shrubs.

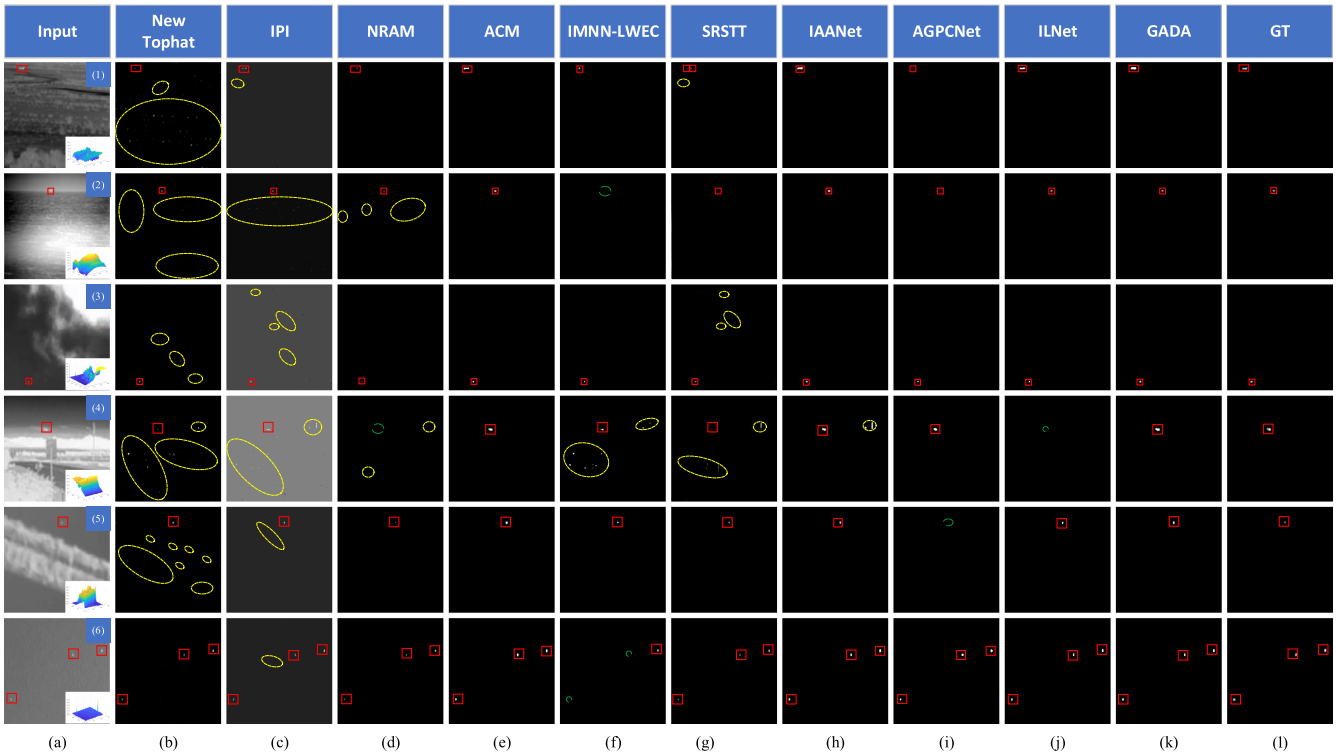

Fig. 3. Detection results obtained on the SIRST dataset using different detection methods. Correctly detected regions, missed alarms, and FAs are highlighted in red, green, and yellow, respectively.

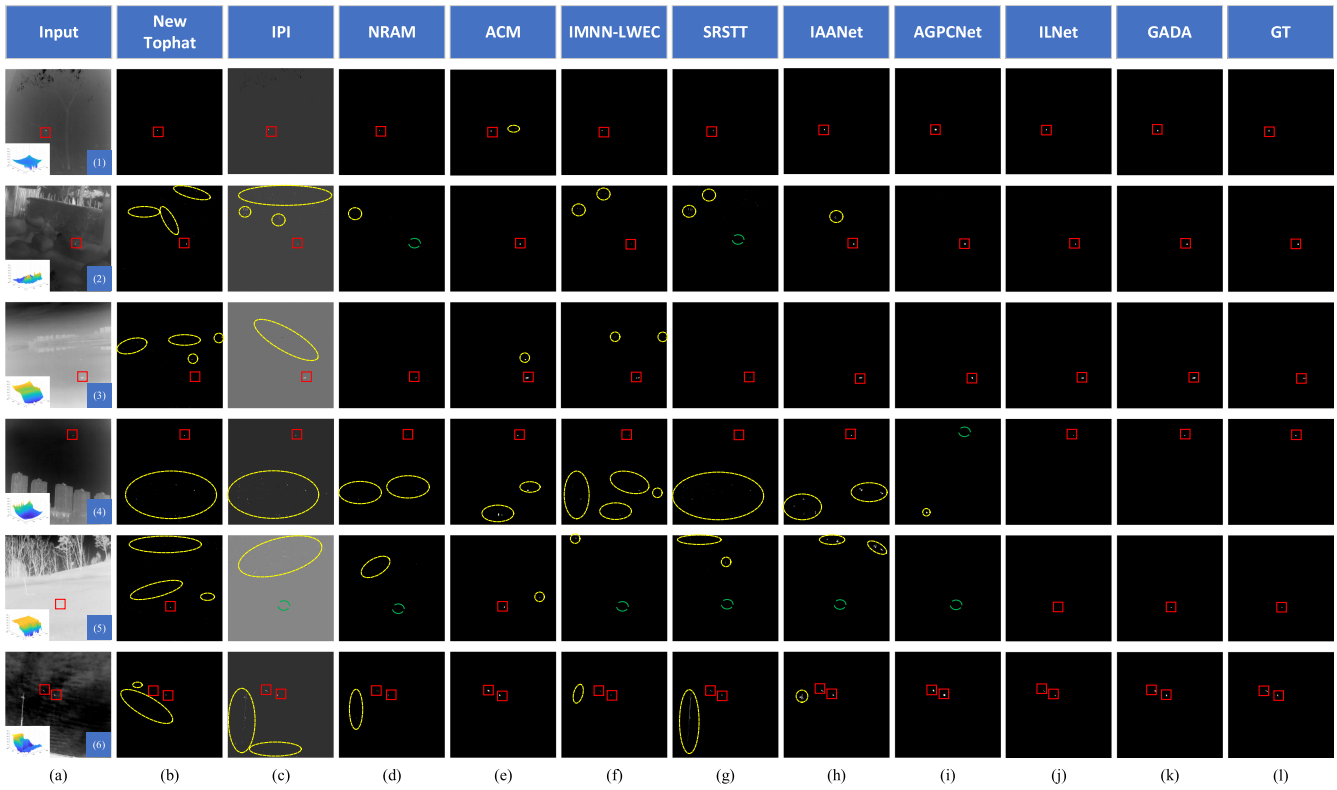

Fig. 4. Detection results obtained on the IRSTD-1 k dataset using different detection methods.

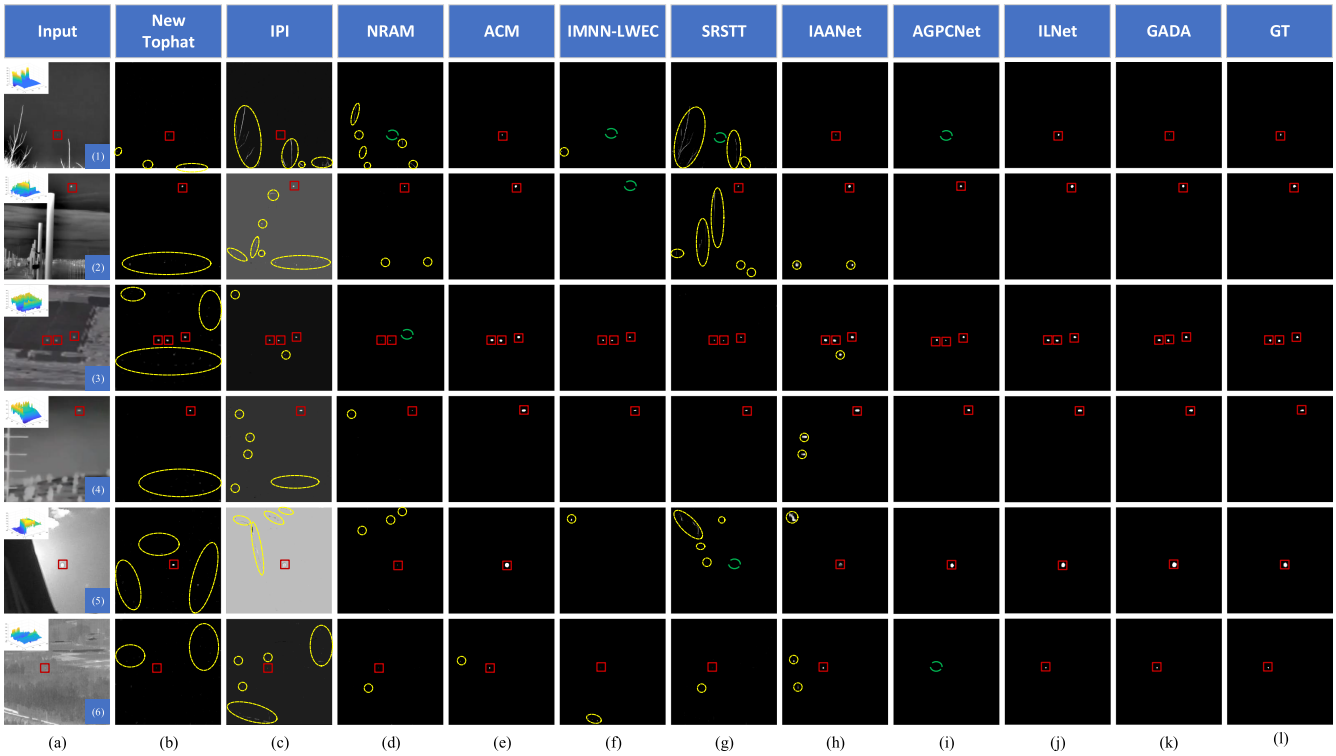

Fig. 5. Detection results obtained on the MSISTD dataset using different detection methods.

TABLE I
$F1$, IoU, AND AUC VALUES OBTAINED BY DIFFERENT STATE-OF-THE-ART METHODS ON THE SIRST AND IRSTD-1K DATASETS.

	SIRST			IRSTD-1k			MSISTD		
	$F1(Precision, Recall)$	IoU	AUC	$F1(Precision, Recall)$	IoU	AUC	$F1(Precision, Recall)$	IoU	AUC
New Tophat	0.52 (0.50, 0.72)	0.41	0.83	0.27 (0.23, 0.63)	0.18	0.66	0.27 (0.23, 0.52)	0.19	0.59
IPI	0.62 (0.90, 0.50)	0.47	0.93	0.42 (0.49, 0.69)	0.32	0.80	0.46 (0.58, 0.54)	0.38	0.70
NRAM	0.70 (0.85, 0.65)	0.59	0.79	0.36 (0.52, 0.36)	0.25	0.59	0.37 (0.56, 0.33)	0.28	0.56
ACM	0.77 (0.91, 0.69)	0.64	0.94	0.65 (0.61, 0.76)	0.51	0.89	0.61 (0.91, 0.48)	0.47	0.93
IMNN-LWEC	0.49 (0.85, 0.38)	0.36	0.74	0.46 (0.57, 0.48)	0.34	0.64	0.31 (0.64, 0.24)	0.22	0.62
SRSTT	0.59 (0.54, 0.76)	0.46	0.90	0.31 (0.31, 0.60)	0.23	0.68	0.34 (0.41, 0.38)	0.22	0.71
IAANet	0.69 (0.90, 0.59)	0.55	0.95	0.65 (0.69, 0.71)	0.53	0.911	0.64 (0.78, 0.61)	0.52	0.95
AGPCNet	0.54 (0.68, 0.51)	0.45	0.85	0.46 (0.48, 0.50)	0.36	0.82	0.50 (0.70, 0.43)	0.40	0.78
ILNet	0.59 (0.54, 0.76)	0.62	0.97	0.60 (0.81, 0.52)	0.48	0.906	0.70 (0.91, 0.61)	0.58	0.90
GADA	0.82 (0.87, 0.82)	0.71	0.95	0.65 (0.68, 0.69)	0.53	0.908	0.71 (0.74, 0.74)	0.59	0.85

For all three, higher values indicate better performance. The best results are indicated in red color. The second results are indicated in green.

The shrub belongs to the highlighted region in the image. New Tophat and IPI are able to detect small targets but there are false detections. NRAM, IMNN-LWEC, SRSTT, and AGPCNet have miss and false detections. Image-2 shows a region of columns, where the columns are the highlighted areas in the image. ACM, ILNet, and GADA can accurately detect the targets. IMNN-LWEC has MDs. The other methods all have false detection in the highlighted areas while detecting small targets. Image-3 shows three people on the playground. New Tophat, IPI, and IAANet all have misdetections. IMNN-ILEC detection is incomplete and there are also missed detections. Image-4 shows the screen next to the high post. There are false detections by New Tophat, IPI, NRAM, and IAANet. All methods can detect small targets, but the detection results of NRAM and SRSTT are not complete. Image-5 shows the sky image with buildings. New Tophat, IPI, NRAM, IMNN-ILEC, SRSTT, and IAANet all have areas of misdetection. There is also a missed detection in SRSTT. The NRAM detection result is incomplete. Image-6 shows the grass by the roadside, the target is located in the grass and the intensity is low. New Tophat, IPI, NRAM, ACM, IMNN-ILEC, SRSTT, and IAANet all have false detection. AGPCNet has missed the detection phenomenon. Also, the detection results of the model-driven based methods are incomplete.

2) Numerical Evaluation: To more accurately demonstrate the validity of our method, we quantitatively analyzed the experimental results using three metrics, $F1$, IoU, and AUC, and then plotted the ROC curves. Table I demonstrates the results of various metrics and shows the comparison results between the three datasets, with the best results marked in red and the second results indicated in green.

As we can see from Table I, data-driven approaches are usually more effective than model-driven approaches. On the three datasets, GADA achieves the best results on $F1$. The values of $F1$ are 0.82, 0.65, and 0.71. It shows that at the pixel level,

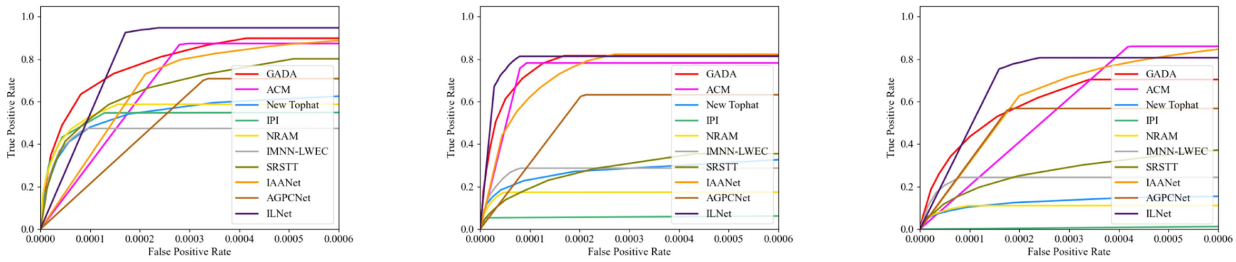

Fig. 6. ROC curves of different methods on different datasets. (a) SIRST. (b) IRSTD-1k. (c) MSISTD.

GADA achieves an optimal balance between false positives and missed detections. Also, we obtained better $Precision$ and $Recall$. Although ACM and ILNet can achieve high $Precision$ on different datasets, the $Recall$ is low, which leads to a lower overall performance of both.

IoU is a metric used to evaluate the localization accuracy. In the experiment, we calculate the average IoU between each real image and predicted image in the test set. In the evaluation results of the three datasets, it can be seen that the IoU values of the traditional detection methods are lower. When the infrared image background is complex with heavy noise and clutter, the traditional detection methods are prone to MDs and FAs, resulting in a decrease in the intersection area and an increase in the concatenation area. The IoU of the traditional methods are all below 0.6 on the SIRST dataset, and below 0.4 on the IRSTD-1 k and MSISTD datasets. The deep learning-based method is better for background suppression. However, most of the detection results show that the location can be detected but the detection is incomplete, which leads to a low value of IoU. The IoU values are 0.71, 0.53, and 0.59 for GADA on three datasets, respectively. It indicates that the predicted images obtained by GADA have a higher degree of overlap with the real image.

The ROC curves are shown in Fig. 6, with the horizontal axis indicating the false positive rate and the vertical axis indicating the true positive rate. The value of AUC on the SIRST dataset is 95%. The value of AUC on the IRSTD-1 k dataset is 91%, which is lower as compared to the SIRST dataset. This is because the dataset contains targets of different intensities and sizes and backgrounds with more noise and clutter. The value of AUC on the MSISTD dataset is 85%. The scenarios in this dataset are more complex and have lower target scale and SCR compared to the other two datasets.

C. Discussion

1) Parametric Analysis: In this part, we validate the setting of the key parameters using experiments. Two parameters α and β are proposed in (9), which represent the weights of reconstruction loss and perception loss, respectively. Our aim is to make the generated background image approximate the original image and not contain small target images as much as possible, so different values of the two parameters are tested. The test results are shown in Table II.

TABLE II
QUALITATIVE RESULTS FOR DIFFERENT VALUES OF α AND β IN (9)

α-β	$Precision$	$Recall$	$F1$	IoU	AUC
1-1	0.8094	0.7561	0.7429	0.6224	0.9356
1-5	0.8643	0.7376	0.7693	0.6601	0.9148
1-10	0.8296	0.6862	0.7244	0.6037	0.8895
5-1	0.8489	0.8298	0.8176	0.7061	0.9513
5-5	0.8723	0.8181	0.8224	0.7123	0.9488
5-10	0.7925	0.8277	0.7787	0.6608	0.9551
10-1	0.8282	0.8096	0.7904	0.6795	0.9206
10-5	0.8362	0.7765	0.7848	0.6729	0.9052
10-10	0.8362	0.7837	0.7824	0.6639	0.9623

The Bold values indicate the best value of each evaluation indicator at different α -β values.

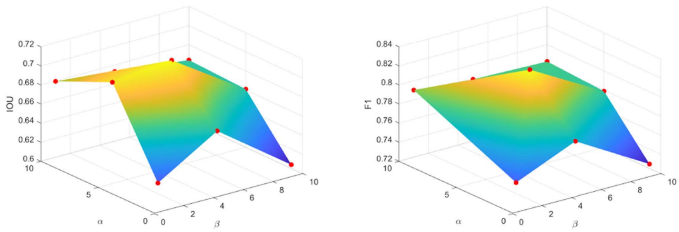

Fig. 7. Three-dimensional demonstration of changes in evaluation metrics IoU and $F1$ for different combinations of parameters α, β. (a) IoU. (b) $F1$.

As can be seen from Table II, the weights of reconstruction loss and perceptual loss are assigned as 1, 5, and 10, respectively. As the assigned values increase sequentially, the test results show an increasing and then decreasing tendency. The evaluation indexes $F1$ and IoU reach 0.8224 and 0.7123 when the values of both α and β are 5. In the test results of other values, the AUC can reach 0.96 when the values of both α and β are 10, but the expressiveness of $F1$ and IoU is much less than that of the results when the values are both 5. At the same time, the AUC can reach 0.95 when the values are both 5.

To visualize the effect of the two parameters on the experimental results more intuitively, 3-D planes were plotted for the values of IoU and $F1$, respectively. As shown in Fig. 7, it can be seen that the values of IoU and $F1$ are located at the vertices of the image when the weight values are both 5.

2) Computational Complexity Analysis: All experiments were implemented on a computer with a 2.50-GHz Intel Core i7-11700 CPU and a Nvidia GeForce 3090GPU. This part has discussed the computational complexity of different methods.

TABLE III
COMPUTATIONAL COMPLEXITY OF DIFFERENT METHODS.

Methods	New Tophat	IPI	NRAM	ACM	IMNN-LWEC
FPS (Hz)	47 (C)	0.2 (C)	0.7 (C)	4.8 (C) 23 (G)	0.4 (C)
Methods	SRSTT	IAANet	AGPCNet	ILNet	GADA
FPS (Hz)	0.07 (C)	0.5 (C) 11 (G)	1.9 (C) 8 (G)	2.3 (C) 13 (G)	1.1 (C) 21 (G)

(C) means the platform is CPU, (G) means the platform is GPU.

TABLE IV
$F1$, IoU, AND AUC VALUES OBTAINED BY DIFFERENT MODELS ON THE SIRST DATASETS

	$Precision$	$Recall$	$F1$	IOU	AUC
FCN	0.7661	0.4173	0.5151	0.3870	0.9445
U-Net	0.8723	0.8181	0.8224	0.7123	0.9488
U-Net++	0.8120	0.7827	0.7542	0.6394	0.9323

FPS is used as an evaluation indicator that indicates the number of images that can be processed per second. The FPS of the model-driven methods is computed on the CPU and the data-driven methods are computed on both CPU and GPU. The experimental results are shown in Table III.

From Table III, it can be seen that New Tophat has the highest FPS value on CPU, but the method is simpler and has more false positives and misses in the detection results. The data-driven model has higher FPS values in ACM, lowest in IAANet, and similar in AGPCNet, ILNet, and GADA. Both ACM and GADA perform better on GPU.

3) Effect of Different Segmentation Models on GADA: GADA uses the differential map for precision extraction to achieve target localization. Therefore, after obtaining the differential map, we need to choose a suitable image segmentation network to complete the target detection task. In this part, the effects of different segmentation models on the experimental results are experimentally verified. Different networks were tested on the SIRST dataset, such as FCN [50], U-Net, and UNet++ [51].

From Table IV, FCN is not sensitive enough to image details, resulting in segmentation results that are not fine enough. It does not take into account the pixel-to-pixel relationship, resulting in a lack of spatial consistency. Therefore, the $Precision$ and $Recall$ of FCN segmentation results are low. The IoU values are also lower due to more FAs in the detection results. UNet++ fuses several different layers of features, and the decoder can feel the feature maps under different horizons. However, due to only fusing the information of the next layer, there is a loss of edge information and position information in the segmentation result. The detection result of UNet++ is greatly improved compared to FCN, but compared to U-Net, the detection result of U-Net is more excellent.

V. CONCLUSION

In this article, we propose the GADA model for infrared small target detection. Unlike existing CNN-based detection methods, we do not directly extract small targets from the background containing a lot of noise and clutter. Instead, the approximate background image and the image with a small target are obtained by generative adversarial differential analysis. Then the prediction results are obtained by precision extraction. We conducted extensive experiments on three public datasets. Compared with the baseline method, GADA has advantages in terms of $F1$, IoU, and AUC, and it can adapt to the task of detecting small targets in a variety of different scenarios. The discussion section shows that the detection results are better when the weight of the loss function is 5 and the precision extraction network structure is U-Net. In summary, GADA can better achieve the purpose of target detection. In future works, we will improve the structure of the generative adversarial network to make the generated background more closely resemble the original image. To make the differential image focus more on small targets to improve the accuracy of image segmentation, we also aim to improve the performance of the differential analysis.

REFERENCES

[1] S. Huang, Y. Liu, Y. He, T. Zhang, and Z. Peng, "Structure-adaptive clutter suppression for infrared small target detection: Chain-growth filtering," *Remote Sens.*, vol. 12, no. 1, 2020, Art. no. 47.

[2] Z. Cao, X. Kong, Q. Zhu, S. Cao, and Z. Peng, "Infrared dim target detection via mode-k1k2 extension tensor tubal rank under complex ocean environment," *ISPRS J. Photogrammetry Remote Sens.*, vol. 181, pp. 167–190, 2021.

[3] Y. Sun, J. Yang, and W. An, "Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3737–3752, May 2021.

[4] W. Zhang, M. Cong, and L. Wang, "Algorithms for optical weak small targets detection and tracking: Review," in *Proc. Int. Conf. Neural Netw. Signal Process.*, 2003, pp. 643–647.

[5] X. Wang, Z. Peng, P. Zhang, and Y. He, "Infrared small target detection via nonnegativity-constrained variational mode decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1700–1704, Oct. 2017.

[6] Y. Liu and Z. Peng, "Infrared small target detection based on resampling-guided image model," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[7] T. Zhang, L. Li, S. Cao, T. Pu, and Z. Peng, "Attention-guided pyramid context networks for detecting infrared small target under complex background," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 4, pp. 4250–4261, Aug. 2023.

[8] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Opt. Photon.*, vol. 3809, pp. 74–83, 1999.

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

[10] M. Zhao, L. Li, W. Li, R. Tao, L. Li, and W. Zhang, "Infrared small-target detection based on multiple morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 6077–6091, Jul. 2021.

[11] 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.

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

[13] J. Liu, H. Wang, L. Lei, and J. He, "Infrared small target detection utilizing halo structure prior-based local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[14] R. Kou, C. Wang, Q. Fu, Y. Yu, and D. Zhang, "Infrared small target detection based on the improved density peak global search and human visual local contrast mechanism," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6144–6157, 2022.

[15] 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.
[16] Y. He, M. Li, J. Zhang, and Q. An, "Small infrared target detection based on low-rank and sparse representation," *Infrared Phys. Technol.*, vol. 68, pp. 98–109, 2015.
[17] C. Zhang, Y. He, Q. Tang, Z. Chen, and T. Mu, "Infrared small target detection via interpatch correlation enhancement and joint local visual saliency prior," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
[18] M. Zhang, R. Zhang, Y. Yang, H. Bai, J. Zhang, and J. Guo, "ISNet: Shape matters for infrared small target detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 867–876.
[19] H. Wang, L. Zhou, and L. Wang, "Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 8508–8517.
[20] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Asymmetric contextual modulation for infrared small target detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2021, pp. 949–958.
[21] B. Li et al., "Dense nested attention network for infrared small target detection," *IEEE Trans. Image Process.*, vol. 32, pp. 1745–1758, 2023.
[22] K. Wang, S. Du, C. Liu, and Z. Cao, "Interior attention-aware network for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022.
[23] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.
[24] B. Zhao, C. Wang, Q. Fu, and Z. Han, "A novel pattern for infrared small target detection with generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4481–4492, May 2021.
[25] A. C. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, pp. 23–34, 2004.
[26] L. Khelifi and M. Mignotte, "Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis," *IEEE Access*, vol. 8, pp. 126385–126400, 2020.
[27] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.
[28] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.
[29] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.
[30] P. Du, S. Liu, P. Gamba, K. Tan, and J. Xia, "Fusion of difference images for change detection over urban areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1076–1086, Aug. 2012.
[31] F. Samadi, G. Akbarizadeh, and H. Kaabi, "Change detection in SAR images using deep belief network: A new training approach based on morphological images," *IET Image Process.*, vol. 13, no. 12, pp. 2255–2264, 2019.
[32] Y. Shu, W. Li, M. Yang, P. Cheng, and S. Han, "Patch-based change detection method for SAR images with label updating strategy," *Remote Sens.*, vol. 13, no. 7, 2021, Art. no. 1236.
[33] W. G. C. Bandara and V. M. Patel, "A transformer-based Siamese network for change detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 207–210.
[34] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–35, Aug. 2019.
[35] R. Kou et al., "Infrared small target segmentation networks: A survey," *Pattern Recognit.*, vol. 143, 2023, Art. no. 109788.
[36] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "Learning a discriminative feature network for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1857–1866.
[37] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
[38] Z. Zuo et al., "AFFPN: Attention fusion feature pyramid network for small infrared target detection," *Remote Sens.*, vol. 14, no. 14, 2022, Art. no. 3412.
[39] C. Yu et al., "Infrared small target detection based on multiscale local contrast learning networks," *Infrared Phys. Technol.*, vol. 123, 2022, Art. no. 104107.
[40] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.
[41] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. Fourth Int. Conf. 3D Vis.*, 2016, pp. 565–571.
[42] A. Wang, W. Li, Z. Huang, X. Wu, F. Jie, and R. Tao, "Prior-guided data augmentation for infrared small target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 10027–10040, 2022.
[43] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint l2,1 norm," *Remote Sens.*, vol. 10, no. 11, 2018, Art. no. 1821.
[44] Y. Luo, X. Li, S. Chen, C. Xia, and L. Zhao, "IMNN-LWEC: A novel infrared small target detection based on spatial–temporal tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–22, 2022.
[45] J. Li, P. Zhang, L. Zhang, and Z. Zhang, "Sparse regularization-based spatial–temporal twist tensor model for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–17, 2023.
[46] H. Li, J. Yang, R. Wang, and Y. Xu, "ILNet: Low-level matters for salient infrared small target detection," 2023, *arXiv:2309.13646*.
[47] F. Chen et al., "Local patch network with global attention for infrared small target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 3979–3991, May 2022.
[48] S. Kim, K.-T. Kim, and S. Kim, "Infrared small target discrimination using sequential forward feature selection with AUC mettric," in *Proc. IEEE Int. Conf. Ind. Technol.*, 2014, pp. 641–644.
[49] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
[50] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.
[51] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested U-Net architecture for medical image segmentation," in *Proc. Int. Workshop Deep Learn. Med. Image Anal. Int. Workshop Multimodal Learn. Clin. Decis. Support*, 2018, pp. 3–11.

Zongfang Ma received the B.S. degree in communication engineering and M.S. degree in technology for computer applications from Xi'an University of Architecture and Technology, Xi'an, China, in 2002 and 2006, respectively, and the Ph.D. degree in control science and engineering from Northwestern Polytechnical University, Xi'an, China, in 2011.

He is a Professor with the School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, China. His research interests include intelligent information processing and industrial application of machine vision.

Shuo Pang received the B.S. degree in computer science and technology from China University of Mining and Technology, Jiangsu, China, in 2022. She is currently working toward the M.S. degree in computer science and technology with the College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, China.

Her research interests include target detection and infrared image processing.

Fan Hao received the B.S. degree in electronic information science and M.S. degree in signal and information processing from Xi'an University of Architecture and Technology, Xi'an, China, in 2020 and 2023, respectively. He is currently working toward the Ph.D. degree in integrated circuit science and engineering with the School of Integrated Circuits, Beijing University of Posts and Telecommunications Beijing, China.

His research interests include pattern recognition, infrared image processing, and remote sensing image change detection.