Infrared Maritime Target Detection Based on Iterative Corner and Edge Weights in Tensor Decomposition

Enzhong Zhao , Lili Dong , and Jin Shi

Abstract—High-precision infrared maritime target detection plays an important role in early warning, monitoring, search, and rescue. The methods of decomposing the original image into low-rank background components and sparse target components show favorable detection performance. However, the strong edge interference is also sparse and may be mistakenly taken as target components, resulting in a large number of false alarms and reducing detection accuracy. To solve the problem, we propose an iterative corner and edge weights method based on tensor decomposition. The original image is decomposed into the background component, target component, and additional strong edge interference component. The corner strength is designed as the weight of the target component, and the edge strength is designed as the weight of the interference component in order to separate the target and interference more accurately. And the two weights are designed to be updated during each iteration of the model solution to reduce the impact of initial imprecise weights on detection results. Compared with 8 advanced baseline methods in 10 datasets, the proposed method demonstrates outstanding results and shows engineering application prospects.

Index Terms—Infrared small target detection, interference suppression, iterative weight, tensor decomposition.

I. INTRODUCTION

AS AN important technology in remote sensing [1], [2], [3], infrared maritime target detection plays a key role in the military, rescue, and other fields [4], [5]. High detection accuracy is conducive to the efficient implementation of the task; however, the infrared maritime image may have interference factors similar to the characteristics of the target, such as islands, strong waves, clouds, etc., which are prone to be wrongly detected and affect the detection accuracy [6]. In recent years, many scholars have improved detection performance through different theories.

A. Related Work

Target detection methods can be divided into multiframe methods using spatiotemporal information and single-frame methods using only spatial information. Multiframe methods have favorable applications in static scenes with low real-time requirements. Typical methods include Markov random field [7], pipeline filtering [8], particle filter [9], dynamic programming [10], temporal profile [11], etc. These methods are usually time-consuming because they need to process multiple frames. In contrast, single-frame methods are usually more efficient but use relatively less information, which can be summarized into filter-based methods, local-information-based methods, graph-theory-based methods, deep-learning-based methods, matrix-decomposition-based methods, etc.

Filter-based methods take advantage of the consistency of the image background. The targets in the background that break the consistency are separated based on spatial or frequency domain. The spatial-based methods include: least-mean-square filtering [12], bilateral filter [13], and morphology-based top-hat transform [14]. Besides, the frequency-based methods include high-pass filter [15], wavelet transform [16], etc. Although these methods are less time-consuming, the high-frequency interference is hard to be removed [17], so they are more suitable for image preliminary processing.

Contrast is a significant factor in the human visual system [18], and some researchers use the local information in the image to extract the target. The local contrast method (LCM) [18] is a classical method, and a large number of improvements have been derived based on LCM, including relative LCM [19] for multiscale target detection. Adaptive scale patch-based contrast measure [20] for detection in the complex background and strong noise. The halo-structure-prior-based LCM [21] combines the structure tensor for better performance. Besides, the visual attention model [22], [23] uses the difference-of-Gaussians filter to obtain the saliency map to extract the target. Furthermore, based on the image entropy theory, the local entropy around the target changes because the texture is destroyed, which can be used to extract the target [24], [25]. These methods are susceptible to high-intensity noise.

In recent years, methods based on graph theory have been applied to target detection. The facet kernel and random walker (FKRW) method [26] combines the local contrast with the random walk to separate the target from the background. The study in [27] utilizes the method based on region growth to detect infrared small targets. Besides, the morphological profile methods based on the max-tree and min-tree theories are applied to detect targets with different sizes and intensities [28], [29].

The research of deep learning has attracted more and more attention, which has a good effect on detecting targets of various

The authors are with the Information and Communication Engineering Department, School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China (e-mail: enzhong@dlmu.edu.cn; donglili@dlmu.edu.cn; sj542209@dlmu.edu.cn).

scales and shapes. Deep-learning-based methods do not require human design features but automatically extract abstract features through the designed model. Representative methods include region convolutional neural networks (R-CNNs) [30], fast R-CNNs [31], you only look once [32], transformer networks [33], [34], etc. Recently, the local patch network [35] and robust infrared small target detection network (RISTDnet) [36] combine the local information to obtain better results for infrared small targets. However, the current scarce data samples are difficult to ensure the accuracy of detection temporarily [6].

The method based on matrix decomposition is originally used to recover low-rank data from data contaminated with noise [37]. This kind of method is suitable for small target detection and has wide applicability as it assumes that the target is sparse and not constrained by the shape of the target [38]. The infrared patch-image (IPI) [39] model transforms the original infrared image into a local patch model for the first time. However, the detection effect and effectiveness of IPI still need to be improved [40]. The weighted IPI (WIPI) model [41] is designed to provide different weights for each patch to suppress strong background edges. The total variation regularization and principal component pursuit [42] model solves the nonsmooth and nonuniform images by attaching additional total variation regularization to the background. The study in [38] analyzes the reasons for the undesirable performance of the IPI model with complex backgrounds and proves that more accurate results are obtained by replacing the nuclear norm with the partial sum of singular values. The nonconvex rank approximation minimization (NRAM) model [40] introduces $l_{2,1}$-norm to suppress strong edge interference. The nonconvex optimization with l_p-norm constraint (NOLC) [43] further suppresses the background and improves efficiency by constraining the sparse component using l_p-norm. The target-aware nonlocal low-rank modeling with saliency filtering regularization method [44] introduces a smooth but nonconvex surrogate of the rank and designs a sparse constraint operator based on local entropy. Besides, the Laplacian function is introduced to replace the nuclear norm in [45].

In order to make better use of the nonlocal spatial information of the image and improve the solving speed, the reweighted infrared patch tensor [46] model converts the 2-D matrix model into a 3-D tensor model. Subsequently, the partial sum of the singular value is applied to estimate the tensor rank, and the alternating direction method of multipliers (ADMM) [47] is applied to solve the model in [48]. The Laplacian function is also introduced in [49] to estimate tensor rank and the $l_{1,1,2}$-norm is introduced to measure strong edge interference. The tensor-fibered nuclear norm based on the Log operator (LogTFNN) is introduced in [50] to estimate tensor rank and applied hypertotal variation to tackle complex backgrounds. Besides, the interpatch correlation enhancement method [51] combines the local visual saliency and improves the running speed.

B. Motivation

Tensor-based methods show decent results. However, interference with strong edges such as islands, waves, and clouds in images also has sparse properties. It is necessary to separate these sparse interference components from the target.

In order to converge quickly and detect accurately, this kind of method often needs to calculate the weight of the original image, which may lead to severe false alarms of results when the initial weight introduces interference components.

Based on the above considerations, we propose an infrared maritime target detection method based on iterative corner and edge weights (ICEW). The contributions of this article are as follows.

1) The interference component based on edge strength weight is designed and measured by the $l_{1,1,2}$-norm in the tensor decomposition model to separate interference such as strong edge and wave clutters more accurately.
2) A method to obtain the weights of the target component and interference component in the iterative process is proposed, which avoids the problem that interference is easily introduced into the target component when the weights are obtained only by the original image.

II. PROPOSED METHOD

A. Proposed Model

The IPT model divides the original image into several subcomponents after converting it into a patch tensor. The infrared maritime background component can be regarded as continuous and highly correlated, in other words, the background tensor has low-rank properties. Meanwhile, the small targets only account for a minority of pixels of the image, so the target tensor is sparse. To estimate the rank of the background tensor, we introduce an efficient tensor rank estimation method: The partial sum of the tubal nuclear norm (PSTNN) [52], which protects large singular values with important information during the optimization process. The l_1-norm is used to measure the sparsity of the target components. Besides, islands and waves in the infrared maritime image have sparse properties similar to the target. The strong edge interference seriously affects the detection results. To solve this problem, the $l_{1,1,2}$-norm is introduced, which has been proven to be able to characterize strong edge interference [40], [49], and weights are designed for the target component and the edge interference component, respectively. Based on the above considerations, a principal component analysis model is proposed as follows:

$$\min_{\mathcal{B},\mathcal{T},\mathcal{S}} \|\mathcal{B}\|_{\text{PSTNN}} + \lambda \|\mathcal{T} \odot \mathcal{W}_t\|_1 + \beta \|\mathcal{S} \odot \mathcal{W}_s\|_{1,1,2}$$

$$\text{s.t. } \mathcal{D} = \mathcal{B} + \mathcal{T} + \mathcal{S} \quad (1)$$

where $\mathcal{B}$ denotes the background tensor, $\mathcal{T}$ denotes the target tensor, $\mathcal{S}$ denotes the sparse interference components used to separate the strong edge interference, $\mathcal{D}$ denotes the patch-tensor constructed from the original image. The calculation of $\|\mathcal{B}\|_{\text{PSTNN}}$ is introduced in Definition 1. $\mathcal{W}_t$ represents the weight of the target component; $\mathcal{W}_s$ represents the weight of $\mathcal{S}$; λ and β are compromising parameters. The calculation of $\|\cdot\|_{1,1,2}$ is introduced in Definition 2.

B. ICEWs of Target Component and Interference Component

For faster convergence and accurate separation of the target components, weights are usually introduced for the target components. Here we introduce an efficient structure-tensor-based method that is well-suited for enhancing small targets. This method utilizes the input image to construct the structure tensor and obtain two eigenvalue matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$. The specific processes of constructing structure tensor and obtaining $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are described in detail in Definition 3. Let the corresponding elements in $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ be λ_1 and λ_2. When $\lambda_1 \approx \lambda_2 \approx 0$, the position corresponds to a flat region. When $\lambda_1 \geq \lambda_2 \gg 0$, the position corresponds to a corner region. When $\lambda_1 \gg \lambda_2 \approx 0$, the position corresponds to an edge region [48].

Then, the corner strength weight of the target component based on the structure tensor $\boldsymbol{W}_{cs}$ can be calculated by the Hadamard product between the harmonic mean and the maximum of the eigenvalues $(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$ as

$$\boldsymbol{W}_{cs} = \frac{\det(\mathbf{ST})}{\operatorname{tr}(\mathbf{ST})} \odot \max(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$$

$$= \frac{\mathbf{\Lambda}_1 \odot \mathbf{\Lambda}_2}{\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2} \odot \max(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2) \quad (2)$$

where $\mathbf{ST}$ indicates the structure tensor, the determinant of structure tensor $\det(\mathbf{ST}) = \mathbf{\Lambda}_1 \odot \mathbf{\Lambda}_2$, the trace of structure tensor $\operatorname{tr}(\mathbf{ST}) = \mathbf{\Lambda}_1 + \mathbf{\Lambda}_2$. To get the corner strength weight $\boldsymbol{\mathcal{W}}_{cs}$ of the tensor $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{X}}$ is first reconstructed as a matrix $\boldsymbol{X}$, then the matrix $\boldsymbol{W}_{cs}$ is obtained by (2), and finally, the normalized $\boldsymbol{\mathcal{W}}_{cs}$ is constructed as a tensor. We abbreviate the above process as

$$\boldsymbol{\mathcal{W}}_{cs} = \mathfrak{W}_t(\boldsymbol{\mathcal{X}}). \quad (3)$$

We select four typical scenes, as shown in the first column in Fig. 1. $\boldsymbol{W}_{cs}$ obtained by (2) is shown in the second column. It can be found that the target area is enhanced. However, parts of islands and waves still remain, which may lead to residual noise in the target component and affect the final detection result. Therefore, it is unreasonable to calculate the weight only through the original image. Considering that the strong edge interference in the target and background components will be gradually separated into the interference component during the iteration process, it is theoretically more reasonable to use the superposition of the target component and background component obtained after each iteration to reobtain the weight of the target component.

Based on the above considerations, we propose an iterative weight method, and its comparison with the general method is shown in Fig. 2. Since the weight of the general method is constant, the interference may be introduced in the results eventually. However, the proposed iterative weight method uses the new $\boldsymbol{B} + \boldsymbol{T}$ to calculate the required weights for the next iteration each time, which separates the strong edge interference with sparsity from the target.

Let k be the number of iterations in the model-solving process. Based on the iterative model, the iterative weight of the target component after k iterations $\boldsymbol{\mathcal{W}}_t^{k+1}$ is designed as

$$\boldsymbol{\mathcal{W}}_t^{k+1} = \boldsymbol{\mathcal{W}}_{sw}^k \odot \boldsymbol{\mathcal{W}}_{rec}^k = \frac{1}{(\boldsymbol{\mathcal{T}}^k + \varepsilon) \odot \mathfrak{W}_t(\boldsymbol{\mathcal{B}}^k + \boldsymbol{\mathcal{T}}^k)} \quad (4)$$
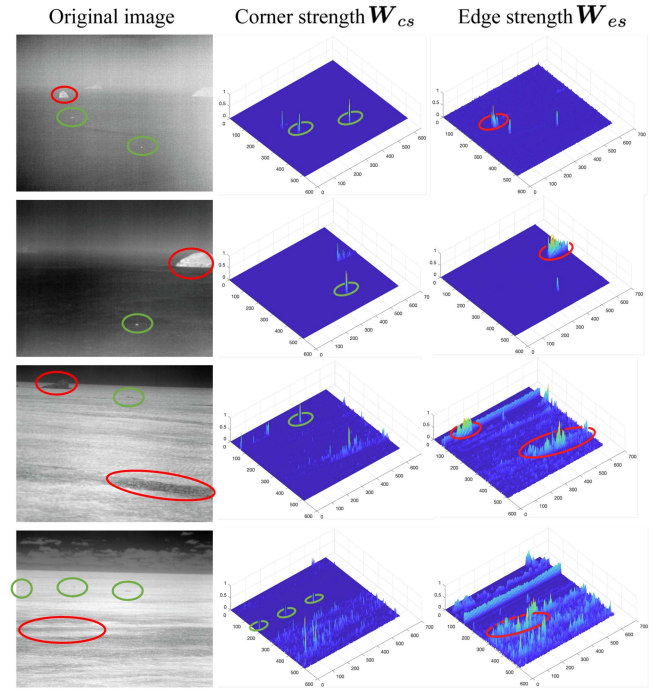

Fig. 1. Weights obtained by the structure tensor. The green circle indicates the location of the target and the red circle indicates the location of the strong edge interference.

where $\boldsymbol{\mathcal{W}}_{sw} = \frac{1}{\boldsymbol{\mathcal{T}} + \varepsilon}$ is a sparsity weight for fast convergence [53], constant ε avoids the denominator being zero. $\boldsymbol{\mathcal{W}}_{rec}$ is the tensor corresponding to elementwise reciprocals of the corresponding elements in $\boldsymbol{\mathcal{W}}_{cs}$.

In order to separate the strong edge interference sufficiently, it is necessary to design a weight for the interference components. Inspired by the corner strength, we combine the relationship of eigenvalues $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ in the structure tensor to design the edge weight of the enhanced edge. The edge strength $\boldsymbol{W}_{es}$ can be calculated as

$$\boldsymbol{W}_{es} = \mathbf{\Lambda}_1 - \mathbf{\Lambda}_2. \quad (5)$$

The third column in Fig. 1 shows the edge strength maps for four typical scenes. The constraint on edge strength is looser, resulting in parts of wave clutters being included in addition to strong edges. The edge interference remaining in corner strength is more prominent in edge strength maps compared to the corner strength maps, which is beneficial to accurately separate the target component and the interference component.

We use the following equation to represent the process of obtaining the normalized edge strength weight $\boldsymbol{\mathcal{W}}_{es}$ of a tensor $\boldsymbol{\mathcal{X}}$ by (5):

$$\boldsymbol{\mathcal{W}}_{es} = \mathfrak{W}_s(\boldsymbol{\mathcal{X}}). \quad (6)$$

Then, the iterative weight of edge component after k iterations $\boldsymbol{\mathcal{W}}_s^{k+1}$ is designed as

$$\boldsymbol{\mathcal{W}}_s^{k+1} = \frac{1}{\mathfrak{W}_s(\boldsymbol{\mathcal{B}}^k + \boldsymbol{\mathcal{T}}^k)}. \quad (7)$$
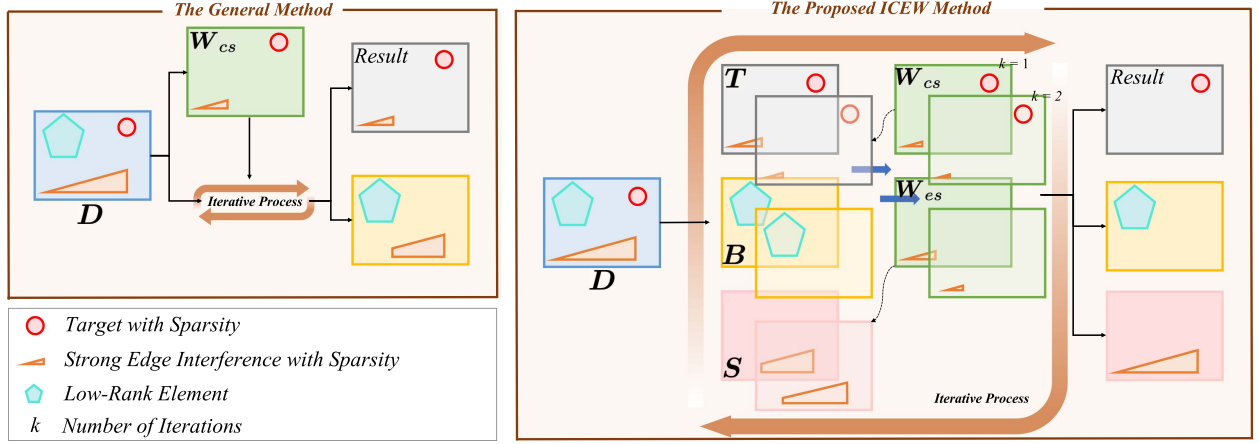
Fig. 2. Schematic of the general model and the proposed ICEW model. The general method on the left obtains the weight outside the iteration process, and the weights of the proposed method on the right are continuously updated during the iteration process. D indicates the input image. W indicates the weight of the component. T indicates the target component. B indicates the background component. S indicates the strong edge component.

C. Solution of the Model

We adopt ADMM to solve the convex optimization problem (1). The augmented Lagrangian function of (1) is shown as follows:

$$\mathcal{L}_\mu = \|\mathcal{X}\|_{\text{PSTNN}} + \lambda \|\mathcal{T} \odot \mathcal{W}_t\|_1 + \beta \|\mathcal{S} \odot \mathcal{W}_s\|_{1,1,2}$$
$$+ < \mathcal{Y}, \mathcal{B} + \mathcal{T} + \mathcal{S} - \mathcal{D} > + \frac{\mu}{2} \|\mathcal{B} + \mathcal{T} + \mathcal{S} - \mathcal{D}\|_F^2 . \tag{8}$$

Then, each component of (8) iterated $k+1$ times can be expressed as follows:

$$\mathcal{B}^{k+1} = \arg\min_{\mathcal{B}} \|\mathcal{B}\|_{\text{PSTNN}}$$
$$+ \frac{\mu^k}{2} \left\| \mathcal{B} + \mathcal{T}^k + \mathcal{S}^k - \mathcal{D} + \frac{\mathcal{Y}^k}{\mu^k} \right\|_F^2 \tag{9}$$

$$\mathcal{T}^{k+1} = \arg\min_{\mathcal{T}} \lambda \left\| \mathcal{T} \odot \mathcal{W}_t^k \right\|_1$$
$$+ \frac{\mu^k}{2} \left\| \mathcal{B}^{k+1} + \mathcal{T} + \mathcal{S}^k - \mathcal{D} + \frac{\mathcal{Y}^k}{\mu^k} \right\|_F^2 \tag{10}$$

$$\mathcal{S}^{k+1} = \arg\min_{\mathcal{S}} \beta \|\mathcal{S} \odot \mathcal{W}_s^k\|_{1,1,2}$$
$$+ \frac{\mu^k}{2} \left\| \mathcal{B}^{k+1} + \mathcal{T}^{k+1} + \mathcal{S} - \mathcal{D} + \frac{\mathcal{Y}^k}{\mu^k} \right\|_F^2 . \tag{11}$$

The subproblem (9) can be solved by partial singular value thresholding (PSVT) operator (see Definition 5) through fast Fourier tensor singular value decomposition (t-SVD) computation (see Definition 6) [48].

The subproblem (10) can be solved by the soft thresholding operator (see Definition 4) [54]

$$\mathcal{T}^{k+1} = S_{\frac{\lambda \mathcal{W}_t^k}{\mu^k}} \left(\mathcal{D} - \mathcal{B}^{k+1} - \mathcal{S}^k - \frac{\mathcal{Y}^k}{\mu^k} \right). \tag{12}$$

Let $\mathcal{Q} = \mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{Y}^k/\mu^k$, the subproblem (11) can be solved by

$$\mathcal{S}^{k+1}(:,:,m) = \mathcal{Q}(:,:,m) \cdot$$
$$\left(1 - \frac{\beta}{\mu^k \|\mathcal{S}(:,:,m)\|_2 \|\mathcal{W}_s(:,:,m)\|_2} \right)_+ \tag{13}$$

where $m = 1, \ldots, n$ indicates the number of tensor patches.

Next, $\mathcal{Y}$ and μ are updated by

$$\mathcal{Y}^{k+1} = \mathcal{Y}^k + \mu^k \left(\mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1} - \mathcal{S}^{k+1} \right) \tag{14}$$
$$\mu^{k+1} = \rho \mu^k \tag{15}$$

where ρ is a constant greater than 1.

D. Polarity Judgment

The polarity of the target indicates the grayscale relationship between the target and the local background. The polarity of targets with grayscales higher than the local background is positive, and the polarity of targets with grayscales lower than the local background is negative. Fig. 3 shows the schematic of the polarity judgment method. After the iterative process converges, the low-rank $\mathcal{B}$ is flatter compared to $\mathcal{S}$. Therefore, in $\mathcal{T}$ calculated by (12), the target whose grayscales are higher than the local background has a positive value, yet the target whose grayscales are lower than the local background has a negative value [55]. However, in certain scenes, the target component may contain sparse interference with opposite polarity, as shown in the matrix T in Fig. 3. Empirically, the target polarity is always consistent in the same scene. Considering that the target is more significant than the interference with the opposite polarity, and the target is also more significant than the interference in $\mathcal{W}_{cs}$, a polarity judgment method based on target-weight similarity is designed as follows:

$$\mathcal{T}^k = \begin{cases} \mathcal{T}_+^k, & \text{if } \|\mathcal{T}_+^k \odot \mathcal{W}_{cs}^k\|_1 \geqslant \|\mathcal{T}_-^k \odot \mathcal{W}_{cs}^k\|_1 \\ -\mathcal{T}_-^k, & \text{otherwise} \end{cases} \tag{16}$$
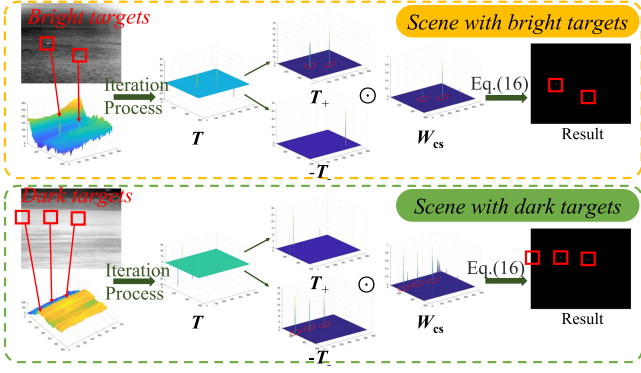
Fig. 3. Schematic diagram of the polarity judgment method. The red circle represents the targets. The final detection result is obtained by comparing the similarity of target components with different polarities (T_+ and T_-) and target weight W_{cs}.

Algorithm 1: ADMM Solver to the Proposed Model.

Input : $\mathcal{T}^0 = \mathcal{B}^0 = \mathcal{S}^0 = \mathcal{Y}^0 = 0, \mathcal{W}_t^0 = \mathcal{W}_s^0 = 1, \mathcal{D}, \lambda, \mu^0, \rho = 1.1, c = 1, k = 0$

1: Fix the others and update $\mathcal{B}^{k+1}$ by PSVT
2: Fix the others and update $\mathcal{T}^{k+1}$ by (12)
3: Fix the others and update $\mathcal{S}^{k+1}$ by (13)
4: Fix the others and update $\mathcal{Y}^{k+1}$ by (14)
5: Fix the others and update $\mathcal{W}_t^{k+1}$ by (4)
6: Fix the others and update $\mathcal{W}_s^{k+1}$ by (7)
7: Update μ by (15)
8: Check the convergence conditions:
   $\frac{\|\mathcal{B}^{k+1}+\mathcal{T}^{k+1}+\mathcal{S}^{k+1}-\mathcal{D}\|_F^2}{\|\mathcal{D}\|_F^2} < \varepsilon$ or $\|\mathcal{T}^{k+1}\|_0 = \|\mathcal{T}^k\|_0$
9: Polarity judgment by (16)
10: Update k: $k = k + 1$
Output : $\mathcal{T}^k$

---

where $\mathcal{T}_+^k = \max(\mathcal{T}^k, 0)$, $\mathcal{T}_-^k = \min(\mathcal{T}^k, 0)$, $\|\mathcal{T}^k \odot \mathcal{W}_{cs}^k\|_1$ indicates the similarity between $\mathcal{T}^k$ and $\mathcal{W}_{cs}^k$. For example, when the target polarity is negative, $-\mathcal{T}_-^k$ and $\mathcal{W}_{cs}^k$ are more similar. Therefore, $-\mathcal{T}_-^k$ is taken as the final result $\mathcal{T}$. The solution procedure for the whole minimization problem is shown in Algorithm 1.

E. Procedure of the Proposed ICEW Method

The flow of ICEW is shown in Fig. 4. The main contributions of this article lie in (4), (7), and (16) in the figure. Different from the general tensor decomposition methods, the weights of the target component and the interference component $\mathcal{W}_t$ and $\mathcal{W}_s$ are obtained by $\mathcal{B} + \mathcal{T}$ each time in the iteration process through (4) and (7), so that the interference in the target is separated more cleanly. Besides, the polarity judgment calculated by (16) removes the interference which is opposite to the target polarity and further ensures the detection accuracy. The following are detailed descriptions of the ICEW process.

First, the original image is constructed as a patch tensor. Second, the original tensor $\mathcal{D}$ is substituted into Algorithm 1 to obtain the $\mathcal{T}$, $\mathcal{B}$, and $\mathcal{S}$ at each iteration. Then, $\mathcal{B}$ and $\mathcal{T}$ are

superimposed and reconstructed into a matrix, and the weights of the $\mathcal{T}$ and $\mathcal{S}$ are obtained by the structure tensor. Then, they are constructed into patch tensors and participated in the next iteration of Algorithm 1. Third, when the iteration is over, $\mathcal{T}$ is divided into positive and negative tensors. After comparing the similarity with $\mathcal{W}_{cs}$, the target polarity can be determined and accurate results obtained. Finally, the retained tensors are reconstructed into matrixes to obtain the final target detection result.

In order to intuitively show the changes of each component and weight in the iterative process, the original image D in Fig. 4 is taken as an example, and reconstruct each tensor into a matrix, as shown in Fig. 5. Each row represents the state of each component and the weight after each iteration. It is obvious that in the first iteration, the distinction between the interference component S and the target component T is not clear. More obvious edge interference exists in the corner strength weight W_{cs}. When after several iterations, the interference in T is gradually eliminated. The interference in W_{cs} is gradually suppressed and more edge interference is introduced in S. After the iteration stops, the similarity between the positive and negative parts of T and the latest W_{cs} is, respectively, compared. Obviously, the polarity of the target T is negative, and the negative part of T is retained and reversed to obtain the final result.

F. Effectiveness of the Proposed ICEW Method

The accuracy of the results benefits from the introduction of the interference component and the update of the weights during each iteration. In order to prove that the proposed method has a significant improvement effect on the detection performance, we set up a comparative experiment. As shown in Fig. 6, where Result$_1$ represents the detection results obtained by neither introducing the interference component nor obtaining weight by iteration in the model. Result$_2$ represents the detection results obtained by iterative corner strength weight without introducing the interference component into the model. Result$_3$ represents the detection results obtained by only introducing the interference component into the model without obtaining iterative weights. The last column shows the final result of ICEW. Numerous edge interference and clutter are introduced in the results of the first model, and the second and third models suppress the interference partly, and the proposed ICEW method performs the best effect.

III. EXPERIMENTS

In order to verify the robustness of the proposed ICEW method, qualitative and quantitative experiments are designed in this section. Common evaluation metrics are used to compare the performance and running time of ICEW with the baseline methods in a variety of scenes.

A. Experimental Setup

1) Dataset: To demonstrate the robustness of ICEW, sequence images of 10 different scenes are provided, all of whose sizes are 640×512 pixels. The typical image in each sequence
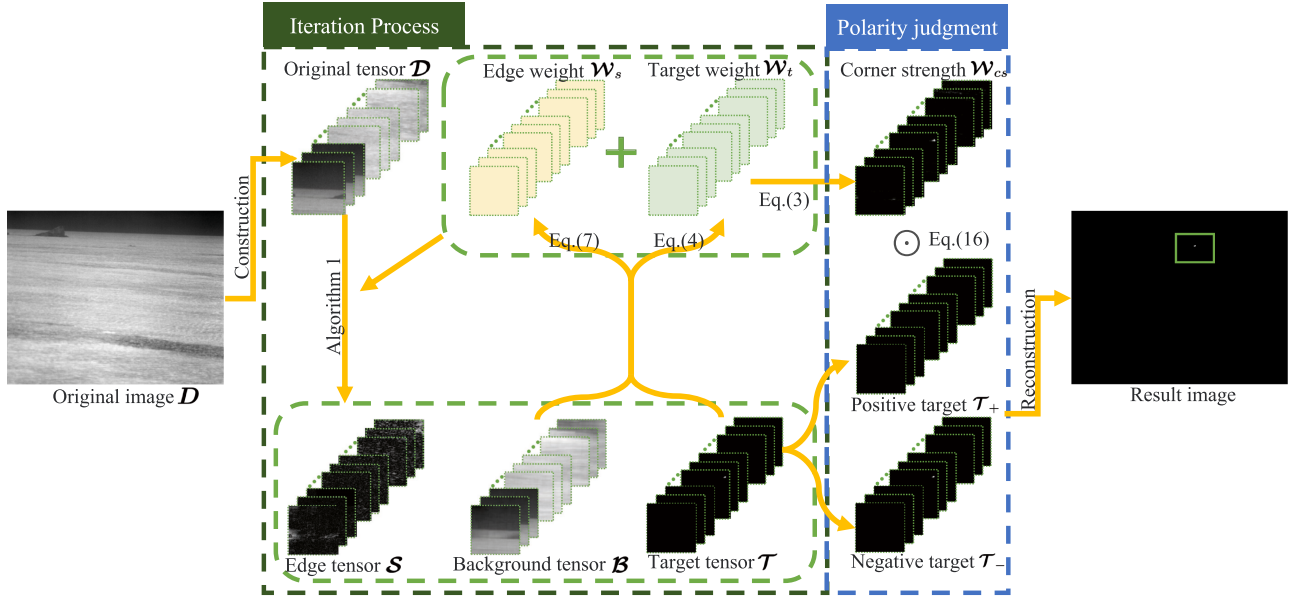
Fig. 4. Flowchart of the proposed ICEW method. The original image is gradually separated into accurate target components, background components, and edge components in the iterative process, and then, the final detection result is obtained by polarity judgment.

TABLE I
DETAIL INFORMATION CONTAINED IN 10 DIFFERENT SCENES

	Target Size	Frames	Interference
(a)	12×6, 13×8, 12×7, 12×7, 31×11	150	Cloud
(b)	9×10, 12×12	51	Backlighting
(c)	13×12, 8×8	87	Wave
(d)	8×8, 8×8	100	Net cage
(e)	11×9, 10×9, 10×8, 19×10, 12×11, 12×10	100	Island
(f)	9×8, 8×8, 9×7, 10×8	107	Island
(g)	16×8, 11×8	150	Island
(h)	14×11	50	Island
(i)	27×8	150	Island&Wave
(j)	17×12, 6×6, 26×8	102	Wave

is shown in Fig. 7. Table I displays the target sizes, the number of frames for each sequence, and the type of interference in each sequence. Considering some methods default the grayscale of the target is higher than their local background, since there are dark targets in datasets (i) and (j), the grayscale of these images is reversed in the experiment for fairness.

2) Evaluation Metrics: Four metrics are selected to evaluate the effectiveness of the proposed method. According to detection accuracy, the detection probability (Pd) and the false alarm rate (Fa) are selected and presented by the receiver operating characteristic (ROC) curve [56], [57]. According to global background suppression ability and local target enhancement ability, the background suppression factor (BSF), and signal-to-clutter ratio gain (SCRG) are selected, respectively. The calculation formula for each parameter is given in the following equations:

$$\text{Pd} = \frac{\text{number of detected true targets}}{\text{number of true targets}} \quad (17)$$

$$\text{Fa} = \frac{\text{number of detected false targets}}{\text{number of detected targets}} \quad (18)$$

$$\text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}} + c} \quad (19)$$

$$\text{SCR} = \frac{|\mu_t - \mu_b|}{\sigma_b + c} \quad (20)$$

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}} + c} \quad (21)$$

where σ_{in} and σ_{out} denote the standard deviation of the original image and final output image excluding the location of the target, respectively. The final output image in this article is the result after normalization and before binarization. μ_t and μ_b denote the average grayscale of the target and background area, respectively. σ_b denotes the standard deviation of the target's local background. SCR$_{\text{in}}$ and SCR$_{\text{out}}$ represent the signal-to-clutter ratio of the original and final output images, respectively. The background-size setting in SCR is taken to be the region obtained by expanding the target region boundary by 20 pixels. Constant c is set to be 0.001 here. According to the definition of the four evaluation metrics, smaller values of Fa indicate more accurate detection, while larger values of Pd, BSF, and SCRG indicate more accurate detection [5].

3) Baseline Methods: Eight baseline methods are selected based on different theories. The name and parameter setting of each method is shown in Table II. Among them, GST [58] is a filter-based method. FKRW [26] is a graph-theory-based method. RLCM [19] is a local-information-based method. NRAM [40], NOLC [43], PSTNN [48], SRWS [59], and LogTFNN [50] are decomposition-based methods. The codes of the proposed ICEW method are available at https://github.com/marny4/ICEW.
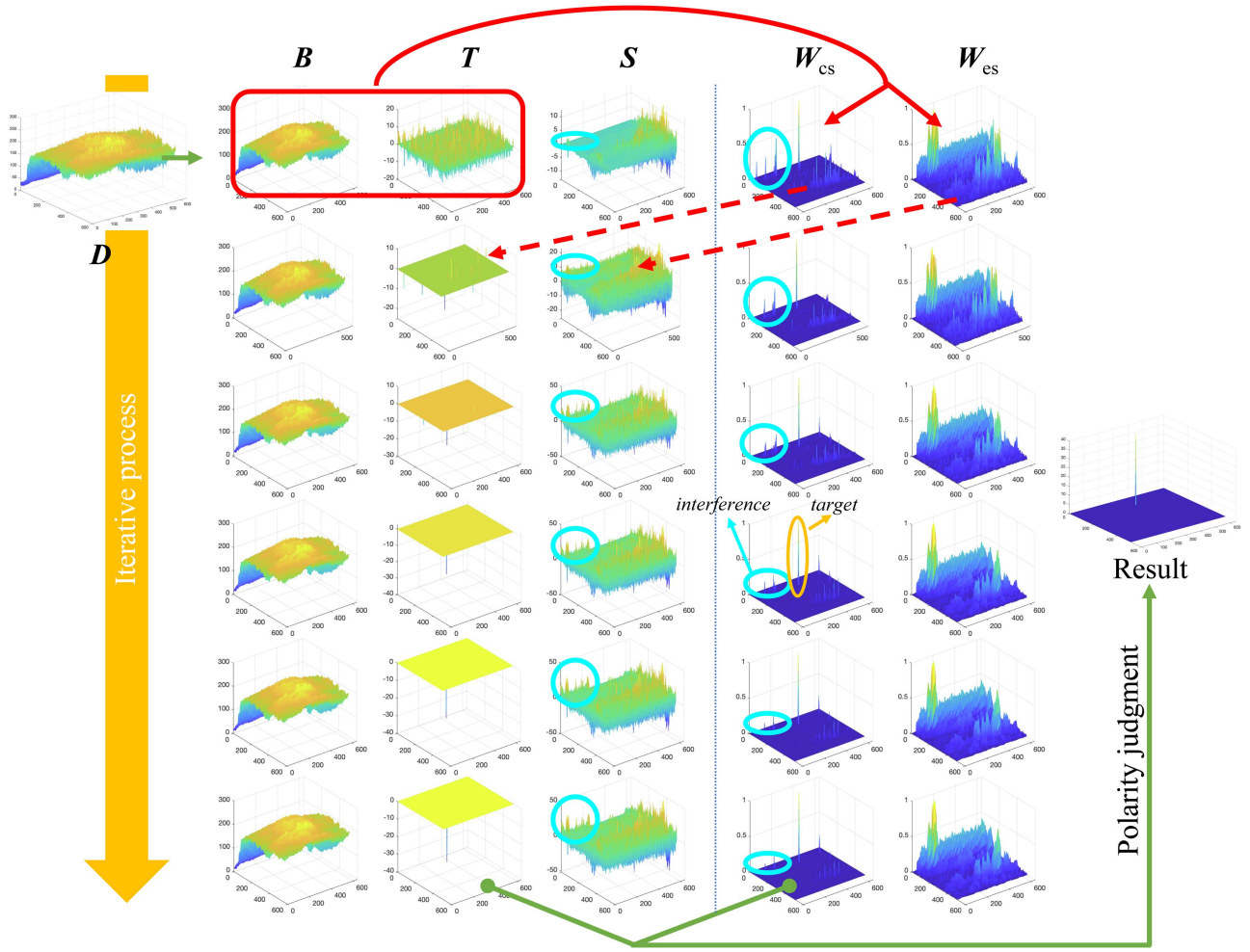
Fig. 5. Variation of each component during the iteration. The cyan circle indicates that some strong edge interference is gradually introduced into the interference component S, and the interference in the corner strength weight W_{cs} of the target component is gradually suppressed during the iteration process.

TABLE II
PARAMETER SETTINGS FOR THE EIGHT BASELINE METHODS

Methods	Parameter settings
GST	$\sigma_1 = 0.6$, $\sigma_2 = 1.1$, boundary width = 5, filter size = 5
FKRW	$k = 4$, $p = 6$, $\beta = 200$, window size: 11×11
RLCM	scale = 3, $K_1 = [2, 5, 9]$, $K_2 = [4, 9, 16]$
NRAM	patch = 50, step = 10, $\lambda = 1/\sqrt{\min(m,n)}$, $\mu = 3\sqrt{\min(m,n)}$, $\gamma = 0.002$, $C = \sqrt{\min(m,n)}/2.5$
NOLC	patch = 30, step = 10, $\lambda = L/\sqrt{\max(size(D))}$, $p = 0.5$
PSTNN	patch = 40, step = 40, $\lambda = \lambda_L/\sqrt{\min(n_1, n_2) * n_3}$
SRWS	patch = 50, step = 50, $\beta = 1/\sqrt{\min(m,n)}$, $\lambda = \lambda_L/\sqrt{\min(m,n)}$, $\gamma = \gamma_L/\sqrt{\min(m,n)}$
LogTFNN	patch = 40, step = 40, $\lambda = 1/\sqrt{\min(n_1, n_2) * n_3}$

B. Analysis of Parameters

By adjusting the key parameters of the proposed ICEW method, the detection effect of four typical scenes (b), (c), (g), and (i) in Fig. 7 are tested. Pd and Fa are used to evaluate the detection effect through the ROC curves so that the performance of the method is in the relatively best state, as shown in Fig. 8. Since the Fa value of the proposed method is low, the range of

Fa is set as [0,0.2]. The curves for the dataset (g) in row 3 rise very fast, which indicates that the target is far more significant than the interference and is not easily missed. When all values of different parameters can achieve high Pd, the values that can achieve a lower Fa can be considered superior.

1) *Patch Size:* The patch size is a parameter related to the target size. Larger patch sizes contribute to the sparsity of the target. Smaller patch sizes contribute to the efficiency of singular value decomposition. The best results are obtained when the patch size is 60 pixels.

2) *Sliding Step:* The sliding step size is also a key parameter that affects the accuracy when constructing the patch tensor. Too small a sliding step leads to a lack of sparsity in the target and increases the time complexity. The best results are obtained when the sliding step is 40 pixels.

3) *Penalty Factor μ:* μ balances the tradeoff among low-rank, sparse, and interference tensors. According to the results in Fig. 8, when $\mu = 1$, the values of Pd are relatively higher and the values of Fa are relatively lower.

4) *Compromising Parameter β:* β is the balance parameter of the interference component, which affects the detection
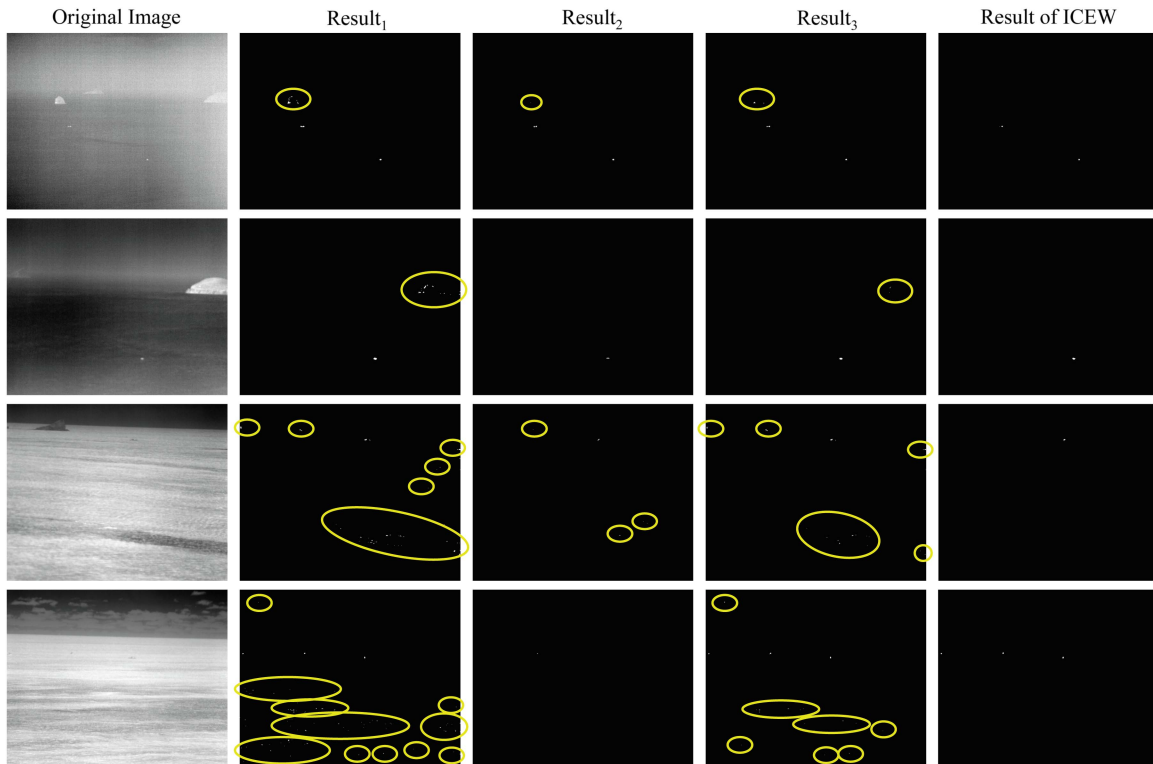
Fig. 6. Comparison results to prove the effectiveness of the proposed ICEW method. The yellow circles indicate the residual noise interference in the results.
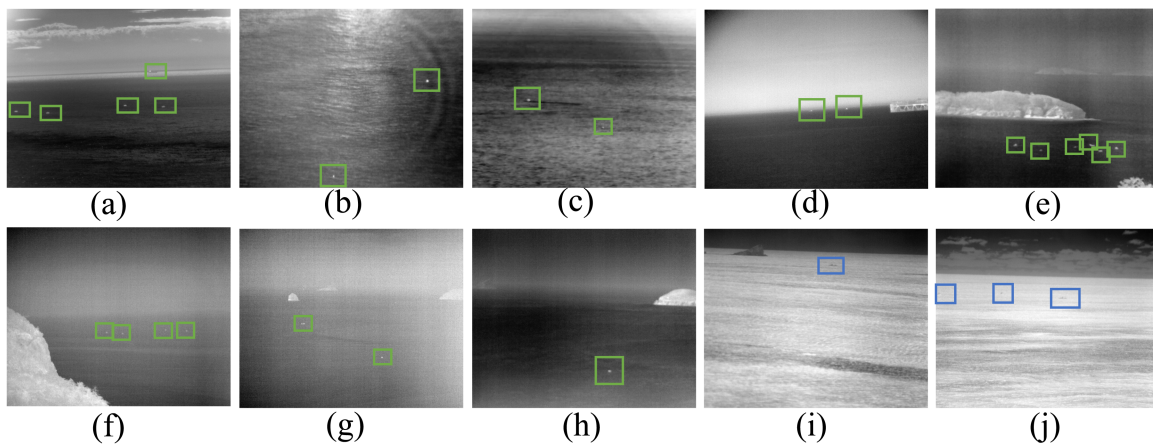

Fig. 7. Dataset in the experiment. (a)–(j) Representative images of each dataset, respectively. The green boxes represent bright targets and the blue boxes represent dark targets.

accuracy. Larger β is prone to introduce more interference into the target component. When the value of β is too small, the target may be introduced into other components. The value of β of 0.2 in this article is reasonable.

5) *Compromising Parameter λ:* λ is the balance parameter of the sparse component, which also affects the detection accuracy. Larger λ may cause the target to be introduced into other components. Smaller λ is prone to introduce more interference into the target component. Referring

to [60], λ is set as $L/\sqrt{\max(p,q)*n}$. L is set to 0.6 in this article.

C. Qualitative Comparison

The detection results of the baselines and the proposed ICEW method in 10 datasets are displayed in Figs. 9–11. The green boxes mark the true targets, and the yellow circles mark the false alarms in all detection results. Among the 8 baselines, GST has
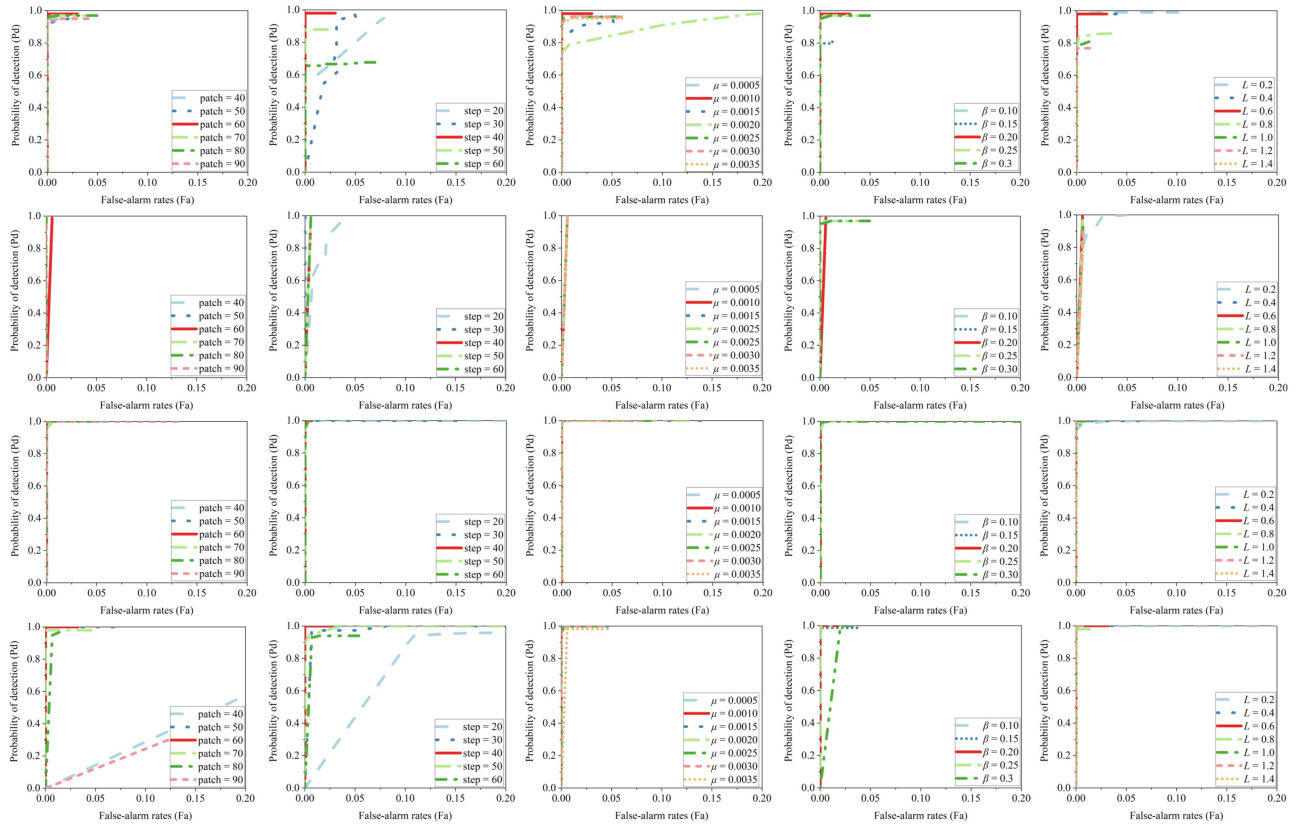
Fig. 8. ROC curves under different key parameters of the proposed ICEW method. Row 1, row 2, row 3, and row 4 indicate the result of the datasets (b), (c), (g), and (i) in Fig. 7, respectively.

few false alarms but has many miss detections, which indicates that GST is difficult to improve the saliency of the target. In addition, GST is unable to suppress strong corner points in Datasets 6 and 8. Although the number of miss detections of FKRW is relatively small, a large number of false alarms caused by clouds, backlighting, islands, and waves remain in the results. The detection accuracy of RLCM is relatively higher, but some weak targets in Datasets 9 and 10 are missed. And because of the design of the sliding window in RLCM, the results lose the morphological characteristics. NRAM is more sensitive to noise, as can be seen in the result of Dataset 7. In addition, NRAM overshrinks the size of the targets, with some false alarms and missed detections in multiple datasets. As a method that introduces a strong edge interference component, NRAM is less capable of suppressing the interference than ICEW. NOLC and SRWS have similar problems as NRAM, but they have relatively few false alarms. The detection rate of PSTNN is relatively high, but it is susceptible to strong edge interference such as islands, which causes a large number of false alarms. Although LogTFNN can detect almost all targets, a large number of false alarms are introduced in the results, which indicates that LogTFNN cannot separate targets and backgrounds well, and a large number of background components are divided into target components. By contrast, ICEW is capable of suppressing noise, waves, islands, and other interference while ensuring a high detection rate, which demonstrates exceptional detection capabilities.

D. Quantitative Comparison

The ROC curves of the proposed ICEW method and the baseline methods under 10 datasets are shown in Fig. 12. Among the 8 baselines, LogTFNN has the worst detection accuracy in most datasets. Although the values of Pd of GST, RLCM, and LogTFNN are relatively high, their corresponding values of Fa are also relatively high, which means that they lack the ability to separate targets from the background. The curves of PSTNN, SRWS, and RLCM rise faster compared with other baselines, which indicates their detection accuracy is relatively higher than other baselines. By contrast, ICEW shows an outstanding effect. The highest values of Pd are obtained with the lowest Fa in almost all datasets compared with baseline methods. Benefiting from our design of iterative weights and additional strong edge interference component, the target components are cleanly separated. Although the values of Pd are not the highest in datasets (b), (c), and (f) compared with some baselines, the values of Fa are extremely low. A small number of missed alarms is more acceptable than massive false alarms.

The results of BSF and SCRG are shown in Tables III and IV. In the tables, bold indicates that ICEW is superior to the baseline methods, and the underlined indicates that ICEW is superior to the other methods combining both BSF and SCRG. For example, the BSF value of ICEW in Dataset 1 is lower than that of GST and SRWS; however, the SCRG value of ICEW in Dataset 1 is
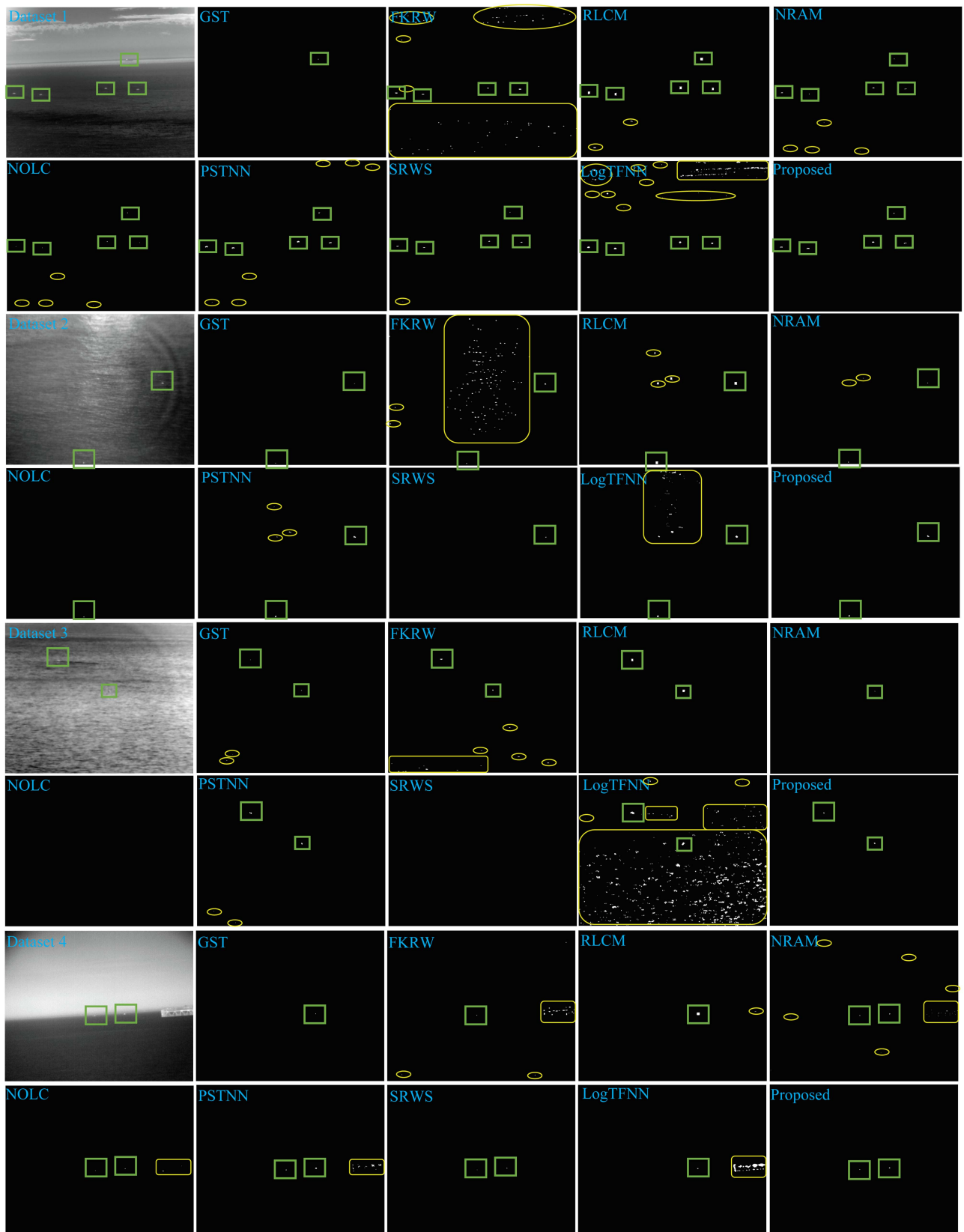
Fig. 9. Qualitative comparison results of Datasets 1 to 4 between the proposed ICEW method and the baseline methods. The yellow circles indicate false alarms.
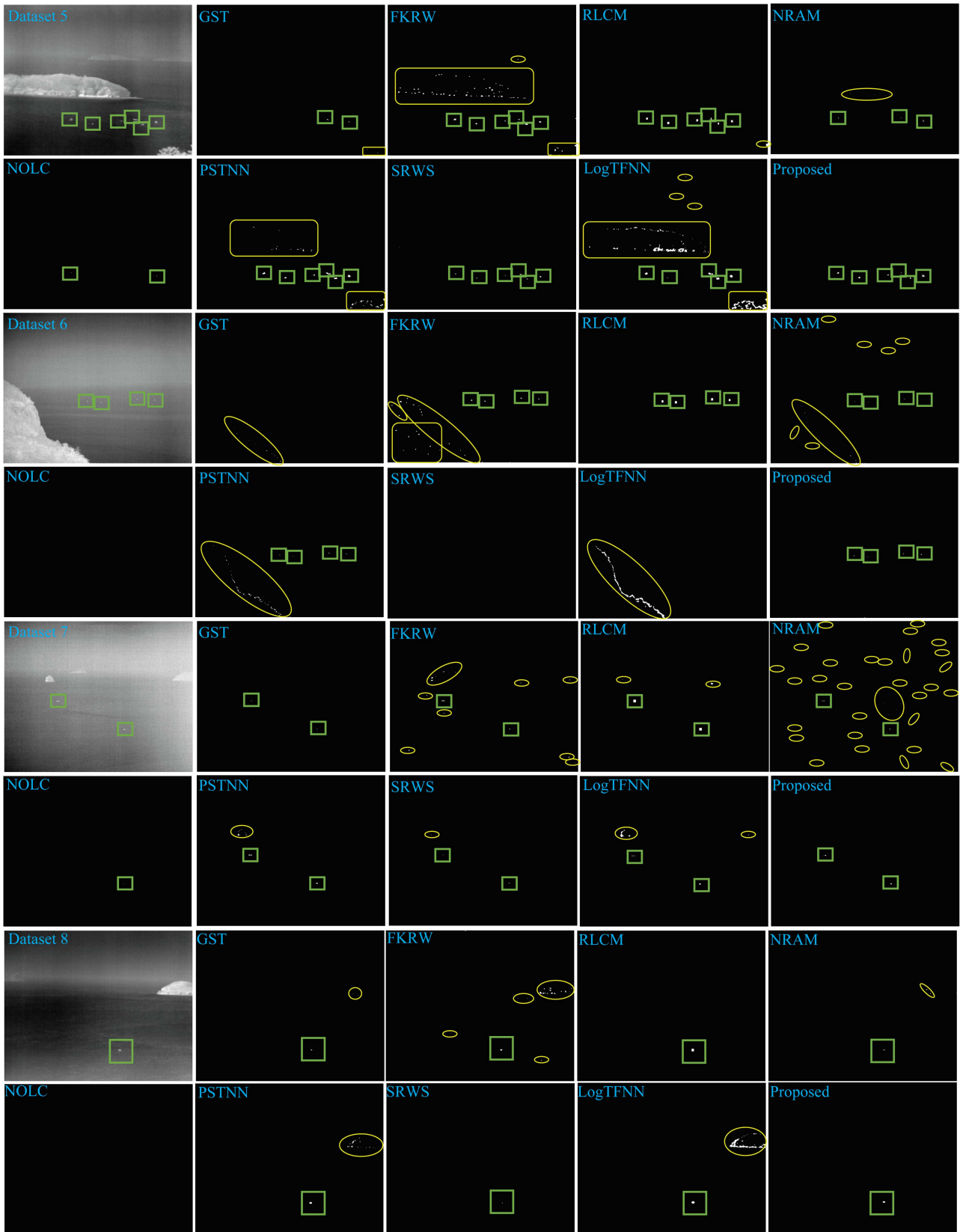
Fig. 10. Qualitative comparison results of Datasets 5–8 between the proposed ICEW method and the baseline methods. The yellow circles indicate false alarms.

Fig. 11.   Qualitative comparison results of Datasets 9–10 between the proposed ICEW method and the baseline methods. The yellow circles indicate false alarms.
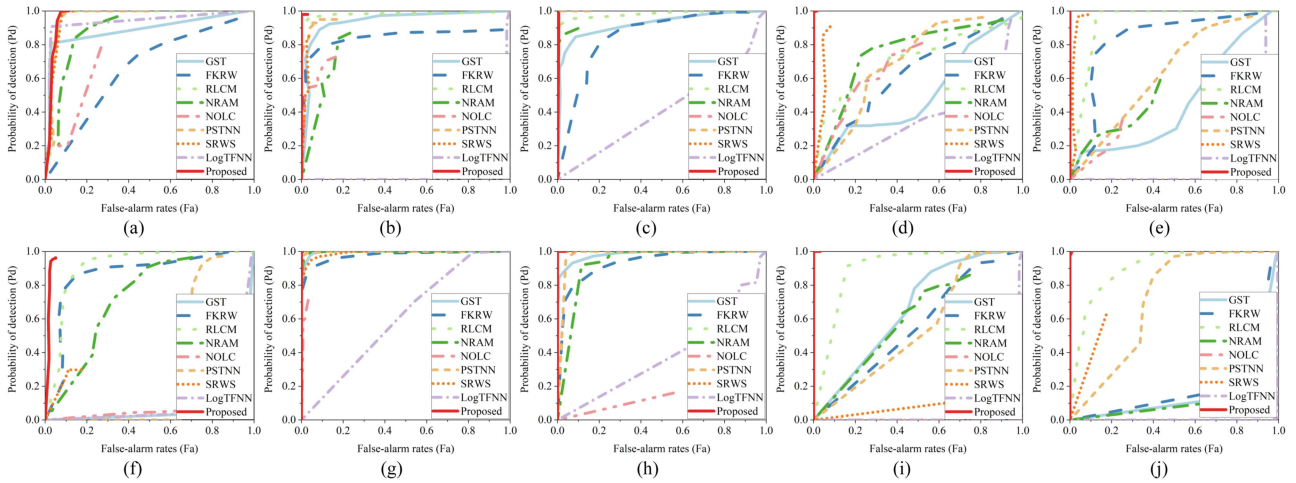

Fig. 12.   ROC curves of the proposed ICEW method and the baseline methods. (a)–(j) Result of 10 different datasets.

obviously higher than that of GST and SRWS, so the underline is introduced for the BSF value of ICEW in Dataset 1.

The BSF values of FKRW, and LogTFNN are relatively low, which indicates that their background suppression abilities are relatively weak and their detection accuracy is more susceptible to complex backgrounds. In most datasets, ICEW demonstrates outstanding background suppression ability, which also shows that ICEW can handle complex scenes better than the baselines,

especially scenes with strong edge interference. Although the BSF values of ICEW in Datasets 1 and 6 are not the highest, their SCRG values are relatively high.

According to the results of SCRG, the signal enhancement performance of GST, NRAM, NOLC, and SRWS is relatively weak, which means their target enhancement performances are not strong and easier to cause miss detections. The higher SCRG value of RLCM profits from the appended adaptive thresholding

TABLE III
AVERAGE BSF OF THE PROPOSED ICEW AND BASELINES

	GST	FKRW	RLCM	NRAM	NOLC	PSTNN	SRWS	LogTFNN	ICEW(Our)
Dataset 1	14.04	4.16	4.60	9.63	10.29	7.66	12.26	0.60	10.58
Dataset 2	13.00	6.23	12.46	12.27	12.99	12.46	13.47	7.27	**33.85**
Dataset 3	5.78	1.82	5.65	5.97	6.21	5.97	6.23	0.52	**8.47**
Dataset 4	17.48	18.14	12.54	16.10	27.11	3.43	35.44	1.51	**41.06**
Dataset 5	6.31	2.05	2.16	2.41	4.79	1.03	6.31	0.25	**6.31**
Dataset 6	8.32	13.43	12.58	11.49	28.64	3.52	29.37	1.23	26.75
Dataset 7	17.10	9.05	11.31	7.96	17.81	4.02	14.30	1.39	**18.08**
Dataset 8	14.82	7.68	12.38	8.68	17.71	3.63	16.03	1.19	**19.99**
Dataset 9	1.74	1.29	7.09	4.72	9.25	1.26	7.20	0.38	9.25
Dataset 10	2.10	1.54	12.56	5.45	12.67	1.38	11.39	0.38	**14.88**

Bold indicates that ICEW is better than the baseline methods.
Underline indicates that the proposed method is better than the baseline methods by combining both BSF and SCRG.

TABLE IV
AVERAGE SCRG OF THE PROPOSED ICEW AND BASELINES

	GST	FKRW	RLCM	NRAM	NOLC	PSTNN	SRWS	LogTFNN	ICEW(Our)
Dataset 1	0.99	11.28	56.02	5.23	2.19	22.69	7.39	22.07	22.70
Dataset 2	1.40	1.96	20.25	2.91	1.97	4.58	1.87	20.26	1.96
Dataset 3	10.81	28.12	121.86	8.21	4.89	46.91	6.01	145.59	47.69
Dataset 4	0.45	2.85	7.73	0.58	0.16	4.66	1.41	4.31	4.56
Dataset 5	42.24	77.52	246.87	145.38	114.93	210.07	70.65	90.55	157.06
Dataset 6	0.20	8.70	43.19	3.95	0.03	5.91	0.15	0.58	9.75
Dataset 7	1.80	3.71	37.34	2.50	0.57	8.74	1.72	4.36	5.81
Dataset 8	1.61	3.39	26.59	2.26	0.38	7.98	1.42	5.52	5.78
Dataset 9	33.56	143.95	12.84	12.49	0	105.67	1.95	91.66	105.78
Dataset 10	28.04	143.93	141.93	8.77	0	146.49	7.30	20.13	92.28

Underline indicates that the ICEW is better than the baseline methods by combining both BSF and SCRG.

TABLE V
AVERAGE RUNTIME OF THE PROPOSED ICEW AND BASELINES ON DATASET 1

	GST	FKRW	RLCM	NRAM	NOLC	PSTNN	SRWS	LogTFNN	ICEW(Our)
Time (s)	0.01	0.16	12.02	19.56	2.33	0.27	1.67	2.87	0.92

at the end of the method. Although the SCRG value of FKRW, PSTNN, and LogTFNN in some datasets are higher than that of ICEW, their BSF values are relatively low, which indicates that the global noise in the results is not well suppressed. Therefore, considering BSF or SCRG alone cannot measure the detection effect well. Considering BSF and SCRG simultaneously, the detection effect of the proposed ICEW method is better than that of the baseline methods. In summary, the proposed ICEW method can separate targets from strong edge interference more accurately compared with the baseline methods.

E. Complexity Analysis and Runtime Comparison

Here, we briefly discuss the complexity of ICEW. The main time consumption comes from the following three aspects: 1) the conversion between tensor and matrix, 2) the tensor singular value decomposition, and 3) fast Fourier transform (FFT) operation. Let the input image be $\mathcal{D} \in \mathbb{R}^{m \times n}$, the patch size be $n_1 \times n_2$, the numbers of patches be n_3, and the iteration times be k. Then, the complexity of the conversion between tensor and matrix is $\mathcal{O}(kmn/(n_1 n_2))$. The complexity of the tensor singular value decomposition is $\mathcal{O}(kn_1 n_2^2[(n_3 + 1)/2])$. The complexity of FFT is $\mathcal{O}(kn_1 n_2 n_3 \log(n_1 n_2))$. Finally, the complexity of ICEW is $\mathcal{O}(k(mn/(n_1 n_2) + n_1 n_2^2[(n_3 + 1)/2] + n_1 n_2 n_3 \log(n_1 n_2)))$. All images are run on a Mac computer (2 GHz quad-core, Intel Core I5 CPU, and 16 GB memory), and the codes are implemented in MATLAB 2022a. The average

runtimes comparison results among the proposed ICEW method and baselines on Dataset 1 are shown in Table V. Among all methods, the runtimes of GST, FKRW, and PSTNN are relatively short. The runtime of RLCM is relatively long because of the traverse of the image. Among methods based on matrix or tensor decomposition, the speed of ICEW is relatively fast. Since the matrix and tensor need to be transformed in each iteration, the running time of ICEW is longer compared with PSTNN.

IV. DISCUSSION

The accuracy of infrared maritime detection still has room for improvement. Affected by the marine environment and the characteristics of infrared cameras, image noise, waves, islands, and clouds easily cause a large number of false alarms. In general matrix or tensor decomposition methods, although a few methods also introduce additional edge interference components, such as NRAM in the baseline methods, no weight is designed for the interference components, resulting in the strong edge cannot be accurately separated from the target. In addition, since the weights have an important influence on the result, if the designs of the weights are not accurate, the detection results are likely to remain false alarms. However, designing more accurate weights increases the complexity of the method, and overly strict weights may reduce the robustness of the method. Considering that the interference in the target and the background component will gradually decrease with iterations in the model-solving

process, the weight of the target and the weight of the edge interference are designed based on the structure tensor through the superposition of the target and background components after each iteration, which is advantageous to separate the interference in the target component to the edge interference component. In addition, we keep the negative values generated during the solution process, and these represent negative polarity targets with grayscales lower than the local background. After the iteration stops, the target polarity is judged by the similarity between the target weight and the targets with different polarities, which helps to screen out interference with opposite polarity to the target. Four parameters, Pd, Fa, BSF, and SCRG, are used as quantitative indicators to evaluate the proposed ICEW method, and 10 sequences with different characteristics (including sea-sky line, noise, light interference, wave, island, cloud, etc.) and 8 baseline methods are selected for a comprehensive comparison. Through qualitative and quantitative analysis, compared with other baselines, ICEW shows more robustness and is suitable for target detection in more scenes.

V. Conclusion

In this article, we proposed a target detection method based on the corner and edge weights in the iterative process. Compared with the general tensor decomposition method, we designed the edge weight and introduced the strong edge interference into the additional component. The weights obtained from the original image were not accurate, so we utilized the updated background and target components to obtain new weights in the iterative process to improve the detection accuracy. The outstanding ability of ICEW to suppress strong edge interference was demonstrated on abundant datasets.

However, as a single-frame method, ICEW still remains some false alarms and missed detections, and the iterative solution process is time-consuming, which is difficult to meet real-time detection requirements. Therefore, in future studies, we will focus on using multiframe information to improve detection accuracy and accelerate our method through parallel processing.

Appendix

Here, we collate some mathematical theories relevant to this article.

Definition 1 (PSTNN [52]): Let a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the PSTNN-norm is defined as follows:

$$\|\mathcal{X}\|_{\text{PSTNN}} = \sum_{i=1}^{n_3} \left\| \overline{X}^{(i)} \right\|_{p=N} \tag{22}$$

$$\|X\|_{p=N} = \sum_{i=p+1}^{\min(n_1, n_2)} \sigma_i(X) \tag{23}$$

where $\overline{X}^{(i)}$ denotes the Fourier transform matrix of the ith frontal slice of $\mathcal{X}$, N is the number of the protected singular

value, $\sigma_i(X)(i = 1, \ldots, \min(n_1, n_2))$ indicates the ith singular value of X.

Definition 2 ($l_{1,1,2}$-norm [49]): Let a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the $l_{1,1,2}$-norm of $\mathcal{X}$ can be calculated by

$$\|\mathcal{S}\|_{1,1,2} = \sum_{i,j} \|\mathcal{S}(i,j,:)\|_F \tag{24}$$

where $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, $\| \cdot \|_F$ indicates the Frobenius norm.

Definition 3 (The Construction of Structure Tensor [58]): Let a matrix $X \in \mathbb{R}^{n_1 \times n_2}$, the structure tensor J_ρ of X can be calculated by

$$J_\rho = K_\rho * (\nabla D_\sigma \otimes \nabla D_\sigma) = \begin{bmatrix} K_\rho * I_x^2 & K_\rho * I_x I_y \\ K_\rho * I_x I_y & K_\rho * I_y^2 \end{bmatrix}$$

$$= \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \tag{25}$$

where K_ρ denotes the Gaussian kernel with variance ρ, D_σ denotes the Gaussian kernel to smooth the input image with variance $\sigma > 0$, $\otimes$ denotes the Kronecker product, $I_x = \frac{\partial D_\sigma}{\partial x}$ and $I_y = \frac{\partial D_\sigma}{\partial y}$ denote the gradient of D_σ. Then, two eigenvalue matrices Λ_1 and Λ_2 of J_ρ can be calculated by

$$\Lambda_1, \Lambda_2 = (J_{11} + J_{22}) \pm \sqrt{(J_{11} - J_{22})^2 + 4J_{12}^2}. \tag{26}$$

Definition 4 (Soft Thresholding Operator [54]): Let $X, Y \in \mathbb{R}^{n_1 \times n_2}$. Define a general l_1-norm minimization problem as

$$\underset{X}{\text{argmin}} \, \tau \|X\|_1 + \frac{1}{2}\|X - Y\|_F^2. \tag{27}$$

Then, (27) can be solved by the soft thresholding operator

$$S_\tau(x) = \text{sign}(x) \times \max(|x| - \tau, 0). \tag{28}$$

Definition 5 (PSVT [61]): Let $X, Y \in \mathbb{R}^{n_1 \times n_2}$, $d = \min(n_1, n_2)$. $Y = Y_1 + Y_2 = U_{Y_1} D_{Y_1} V_{Y_1}^H + U_{Y_2} D_{Y_2} V_{Y_2}^H$. Define a general partial sum of singular values minimization problem as

$$\underset{X}{\text{argmin}} \tau \|X\|_{p=N} + \frac{\beta}{2}\|X - Y\|_F^2. \tag{29}$$

Then, the optimal solution of (29) can be expressed by the PVST operator, which is defined as

$$\mathcal{P}_{N,\tau}(Y) = U_Y(D_{Y_1} + D_\tau[D_{Y_2}])V_Y^H$$

$$= Y_1 + U_{Y_2} S_\tau[D_{Y_2}]V_{Y_2}^H \tag{30}$$

where $\tau = \lambda/\beta$, $D_{Y_1} = \text{diag}(\sigma_1^Y, \ldots, \sigma_N^Y, 0, \ldots, 0)$, and $D_{Y_2} = \text{diag}(0, \ldots, 0, \sigma_{N+1}^Y, \ldots, \sigma_d^\gamma)$.

Definition 6 (Fast Fourier t-SVD [48]): Let $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^{n_1 \times n_2}$, $d = \min(n_1, n_2)$. Define a general PSTNN minimization problem as

$$\underset{\mathcal{X}}{\text{argmin}} \tau \|\mathcal{X}\|_{\text{PSTNN}} + \frac{\beta}{2}\|\mathcal{X} - \mathcal{Y}\|_F^2. \tag{31}$$

Then, (31) can be solved by fast Fourier t-SVD described in Algorithm 2.

Algorithm 2: Fast Fourier t-SVD.

Input : $\mathcal{Y}^k, \tau$

1:   Compute $\overline{\mathcal{Y}}^k = \text{fft}(\mathcal{Y}^k, [\,], 3)$

2:   Compute each frontal slice of $\overline{\mathcal{X}}^{k+1}$ by;

for $i = 1 : [(n_3 + 1)/2]$ **do**

   $(\overline{\mathcal{X}}^{k+1})^{(i)} = \mathcal{P}_{N,\tau}(\overline{\mathcal{Y}}^k)^{(i)}$

end

**for** $i = [(n_3 + 1)/2] + 1 : n_3$ **do**

   $(\overline{\mathcal{X}}^{k+1})^{(i)} = \text{conj}((\overline{\mathcal{X}}^{k+1})^{(n_3-i+2)})$

end

3:   Compute $\mathcal{X} = \text{ifft}(\overline{\mathcal{X}}, [], 3)$

REFERENCES

[1] M. Zhang, W. Li, Y. Zhang, R. Tao, and Q. Du, "Hyperspectral and LiDAR data classification based on structural optimization transmission," *IEEE Trans. Cybern.*, vol. 53, no. 5, pp. 3153–3164, May 2023.

[2] Y. Zhang, W. Li, W. Sun, R. Tao, and Q. Du, "Single-source domain expansion network for cross-scene hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 32, no. 2, pp. 1498–1512, Feb. 2023.

[3] Y. Zhang, M. Zhang, W. Li, S. Wang, and R. Tao, "Language-aware domain generalization network for cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, no. 1, pp. 1–12, Jan. 2023.

[4] D. K. Prasad, D. Rajan, L. Rachmawati, E. Rajabally, and C. Quek, "Video processing from electro-optical sensors for object detection and tracking in a maritime environment: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 1993–2016, Aug. 2017.

[5] M. Zhang, L. Dong, D. Ma, and W. Xu, "Infrared target detection in marine images with heavy waves via local patch similarity," *Infrared Phys. Technol.*, vol. 125, Sep. 2022, Art. no. 104283.

[6] D. Liu, L. Cao, Z. Li, T. Liu, and P. Che, "Infrared small target detection based on flux density and direction diversity in gradientvector field," *IEEE Trans. Geosci. Remote Sens.*, vol. 11, no. 7, pp. 2528–2554, May 2018.

[7] G. Lin and W. Fan, "Unsupervised video object segmentation based on mixture models and saliency detection," *Neural Process. Lett.*, vol. 51, no. 1, pp. 657–674, Aug. 2020.

[8] B. Li, X. Zhiyong, J. Zhang, X. Wang, and X. Fan, "Dim-small target detection based on adaptive pipeline filtering," *Math. Problem Eng.*, vol. 2020, pp. 1–15, May 2020.

[9] J. Pak, "Visual odometry particle filter for improving accuracy of visual object trackers," *Electron. Lett.*, vol. 56, no. 17, pp. 884–887, Aug. 2020.

[10] J. Fu, H. Zhang, W. Luo, and X. Gao, "Dynamic programming ring for point target detection," *Appl. Sci.*, vol. 12, no. 3, Jan. 2022, Art. no. 1151.

[11] X. Liu, L. Li, L. Liu, X. Su, and F. Chen, "Moving dim and small target detection in multiframe infrared sequence with low SCR based on temporal profile similarity," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, no. 4, pp. 1–5, Apr. 2022.

[12] P. Lv, S. Sun, C. Lin, and G. Liu, "Space moving target detection and tracking method in complex background," *Infrared Phys. Technol.*, vol. 91, pp. 107–118, Jun. 2018.

[13] T. Bae and K. Sohng, "Small target detection using bilateral filter based on edge component," *J. Infrared Millimeter Terahertz Waves*, vol. 31, no. 6, pp. 735–743, Mar. 2010.

[14] C. Wang and L. Wang, "Multidirectional ring top-hat transformation for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 14, no. 7, pp. 8077–8088, Aug. 2021.

[15] 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, Aug. 2017.

[16] H. Wang and Y. Xin, "Wavelet-based contourlet transform and Kurtosis map for infrared small target detection in complex background," *Sensors*, vol. 20, no. 3, Jan. 2020, Art. no. 755.

[17] M. Zhang, L. Dong, H. Zheng, and W. Xu, "Infrared maritime small target detection based on edge and local intensity features," *Infrared Phys. Technol.*, vol. 119, Dec. 2021, Art. no. 103940.

[18] C. 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, Mar. 2013.

[19] J. Han et al., "Infrared small target detection based on the weighted strengthened local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 9, pp. 1670–1674, Sep. 2021.

[20] Z. Qiu, Y. Ma, F. Fan, J. Huang, and M. Wu, "Adaptive scale patch-based contrast measure for dim and small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, no. 12, pp. 1–5, Dec. 2022.

[21] 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, no. 3, pp. 1–5, Mar. 2022.

[22] L. Dong, B. Wang, M. Zhao, and W. Xu, "Robust infrared maritime target detection based on visual attention and spatiotemporal filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 3037–3050, Feb. 2017.

[23] Y. Chen, B. Song, D. Wang, and L. Guo, "An effective infrared small target detection method based on the human visual attention," *Infrared Phys. Technol.*, vol. 95, pp. 128–135, Dec. 2018.

[24] Z. Zhou, H. Zhang, Z. Wang, and H. Zheng, "The small target detection based on local directional contrast associated with directional entropy," in *Proc. Int. Conf. Digit. Image Process.*, 2019, vol. 11179, pp. 609–617.

[25] H. Zhang and Z. Zhou, "Small target detection based on automatic ROI extraction and local directional gray&entropy contrast map," *Infrared Phys. Technol.*, vol. 107, Jun. 2020, Art. no. 103290.

[26] Y. Qin, L. Bruzzone, C. Gao, and B. Li, "Infrared small target detection based on facet kernel and random walker," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7104–7118, May 2019.

[27] S. Huang, Z. Peng, Z. Wang, X. Wang, and M. Li, "Infrared small target detection by density peaks searching and maximum-gray region growing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1919–1923, May 2019.

[28] 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, Sep. 2020.

[29] M. Zhao, W. Li, L. Li, P. Ma, Z. Cai, and R. Tao, "Three-order tensor creation and tucker decomposition for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, no. 2, pp. 1–16, Feb. 2021.

[30] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[31] R. Girshick, "Fast R-CNN," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1440–1448.

[32] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[33] K. Han et al., "A survey on vision transformer," *IEEE. Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 87–110, Jan. 2023.

[34] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Proc. Euro. Conf. Comput. Vis.*, 2020, pp. 213–229.

[35] 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, Mar. 2022.

[36] Q. Hou, Z. Wang, F. Tan, Y. Zhao, H. Zheng, and W. Zhang, "RISTDnet: Robust infrared small target detection network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, no. 1, pp. 1–5, Jan. 2021.

[37] M. Zhao, W. Li, L. Li, J. Hu, P. Ma, and R. Tao, "Single-frame infrared small-target detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 87–119, Jun. 2022.

[38] Y. Dai, Y. Wu, Y. Song, and J. Guo, "Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values," *Infrared Phys. Technol.*, vol. 81, pp. 182–194, Mar. 2017.

[39] 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, Sep. 2013.

[40] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm," *Remote Sens.*, vol. 10, no. 11, Nov. 2018, Art. no. 1821.

[41] Y. Dai, Y. Wu, and Y. Song, "Infrared small target and background separation via column-wise weighted robust principal component analysis," *Infrared Phys. Technol.*, vol. 77, pp. 421–430, Jul. 2016.

[42] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image Vis. Comput.*, vol. 63, pp. 1–9, Jul. 2017.

[43] T. Zhang, H. Wu, Y. Liu, L. Peng, C. Yang, and Z. Peng, "Infrared small target detection based on non-convex optimization with lp-norm constraint," *Remote Sens.*, vol. 11, no. 5, Mar. 2019, Art. no. 559.

[44] H. Zhu, H. Ni, S. Liu, G. Xu, and L. Deng, "TNLRS: Target-aware non-local low-rank modeling with saliency filtering regularization for infrared small target detection," *IEEE Trans. Image Process.*, vol. 29, no. 9, pp. 9546–9558, Oct. 2020.

[45] Z. Lu, Z. Huang, Q. Song, K. Bai, and Z. Li, "An enhanced image patch tensor decompostion for infrared small target detection," *Remote Sens.*, vol. 14, no. 23, Nov. 2022, Art. no. 6044.

[46] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 10, no. 8, pp. 3752–3767, May 2017.

[47] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[48] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, Feb. 2019, Art. no. 382.

[49] X. Guan, L. Zhang, S. Huang, and Z. Peng, "Infrared small target detection via non-convex tensor rank surrogate joint local contrast energy," *Remote Sens.*, vol. 12, no. 9, May 2020, Art. no. 1520.

[50] X. Kong, C. Yang, S. Cao, C. Li, and Z. Peng, "Infrared small target detection via nonconvex tensor fibered rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, no. 4, pp. 1–21, Apr. 2021.

[51] 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, no. 11, pp. 1–14, Nov. 2021.

[52] T. Jiang, T. Huang, X. Zhao, and L. Deng, "Multi-dimensional imaging data recovery via minimizing the partial sum of tubal nuclear norm," *J. Comput. Appl. Math.*, vol. 372, Jan. 2020, Art. no. 112680.

[53] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, Oct. 2008.

[54] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l_1-minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, 2008.

[55] E. Zhao, L. Dong, and H. Dai, "Infrared maritime small target detection based on multidirectional uniformity and sparse-weight similarity," *Remote Sens.*, vol. 14, no. 21, Oct. 2022, Art. no. 5492.

[56] J. Gao, Y. Guo, Z. Lin, W. An, and J. Li, "Robust infrared small target detection using multiscale gray and variance difference measures," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 5039–5052, Dec. 2018.

[57] P. Yang, L. Dong, and W. Xu, "Infrared small maritime target detection based on integrated target saliency measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 2369–2386, Jan. 2021.

[58] C. Gao, J. W. Tian, and P. Wang, "Generalised-structure-tensor-based infrared small target detection," *Electron. Lett.*, vol. 44, no. 23, pp. 1349–1351, Nov. 2008.

[59] T. Zhang, Z. Peng, H. Wu, Y. He, C. Li, and C. Yang, "Infrared small target detection via self-regularized weighted sparse model," *Neurocomputing*, vol. 420, pp. 124–148, Oct. 2021.

[60] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 5249–5257.

[61] T. Oh, Y. Tai, J. Bazin, H. Kim, and I. S. Kweon, "Partial sum minimization of singular values in robust PCA: Algorithm and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 4, pp. 744–758, Apr. 2015.
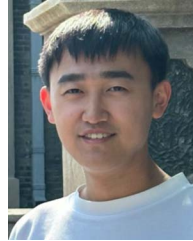
Enzhong Zhao received the B.S. degree in electronic information science and technology and the M.S. degree in electronic science and technology in 2018 and 2021, respectively, from Dalian Maritime University, Dalian, China, where he is currently working toward the Ph.D. degree in information and communication engineering.

His research interests include target detection, image processing, and tunnel lighting.

Lili Dong received the B.S. degree in mechanical design manufacturing and automation, the M.S. degree in machine-electronic engineering, and the Ph.D. degree in instrument science and technology from the Harbin Institute of Technology, Harbin, China, in 2002, 2004, and 2008, respectively.

From 2005 to 2008, she was a Teaching Assistant with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. From 2008 to 2012, she was a Lecturer with Dalian Maritime University and an Associate Professor from 2012 to 2019. Since 2019, she has been a Professor of Information Science and Technology with Dalian Maritime University. She has authored more than 20 articles and 3 inventions. Her research interests include multispectral target recognition, tunnel lighting, and photoelectric detection.

Jin Shi received the bachelor's degree in communication engineering from the Harbin University of Science and Technology, Harbin, China, in 2021. He is currently working toward the master's degree in information and communication engineering with Dalian Maritime University, Dalian, China.

His research interests include target detection, image processing, and electronic image stabilization.