



Article

Image Enhancement-Based Detection with Small Infrared Targets

Shuai Liu ^{1,2}, Pengfei Chen ^{1,2} and Marcin Woźniak ^{3,*}

¹ Key Laboratory of Big Data Research and Application for Basic Education, Hunan Normal University, Changsha 410081, China; liushuai@hunnu.edu.cn (S.L.); 202120293780@hunnu.edu.cn (P.C.)

² College of Educational Science, Hunan Normal University, Changsha 410081, China

³ Faculty of Applied Mathematics, Silesian University of Technology, 44100 Gliwice, Poland

* Correspondence: marcin.wozniak@polsl.pl

Abstract: Today, target detection has an indispensable application in various fields. Infrared small-target detection, as a branch of target detection, can improve the perception capability of autonomous systems, and it has good application prospects in infrared alarm, automatic driving and other fields. There are many well-established algorithms that perform well in infrared small-target detection. Nevertheless, the current algorithms cannot achieve the expected detection effect in complex environments, such as background clutter, noise inundation or very small targets. We have designed an image enhancement-based detection algorithm to solve both problems through detail enhancement and target expansion. This method first enhances the mutation information, detail and edge information of the image and then improves the contrast between the target edge and the adjacent pixels to make the target more prominent. The enhancement improves the robustness of detection with background clutter or noise-flooded scenes. Moreover, bicubic interpolation is used on the input image, and the target pixels are expanded with upsampling, which enhances the detection effectiveness for tiny targets. From the results of qualitative and quantitative experiments, the algorithm proposed in this paper outperforms the existing work on various evaluation indicators.

Keywords: small target detection; infrared small target; autonomous systems; perception capabilities; image enhancement; upsampling



Citation: Liu, S.; Chen, P.; Woźniak, M. Image Enhancement-Based Detection with Small Infrared Targets. *Remote Sens.* **2022**, *14*, 3232. <https://doi.org/10.3390/rs14133232>

Academic Editors: Riccardo Roncella and Mattia Previtali

Received: 16 June 2022

Accepted: 30 June 2022

Published: 5 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autonomous systems refer to systems that can deal with non-programmed or non-preset situations and has certain self-management and self-guidance ability. Compared with automation equipment and systems, autonomous systems can cope with more environments, can complete a wider range of operations and control and have broader application potential. In the future, there are good reasons to believe that autonomy is the ultimate destination of the control field. With recent developments, autonomous systems have successfully achieved ideal results in various key fields such as autopilot and aircraft anti-collision systems.

An infrared small target is a target with low contrast and a signal-to-noise ratio that occupies only a few dozen pixels on the imaging plane when the infrared imaging distance is far. Compared with target detection under visible light, infrared small-target detection has stronger anti-interference ability, stronger night detection ability and more accuracy. This makes the infrared small-target detection technology play an indispensable role in the fields of infrared remote warning, infrared imaging guidance and infrared search and tracking. Therefore, the research on infrared small-target detection technology is of great significance.

One of the key research fields of autonomous system perception is infrared small-target detection. This is because an infrared image has four advantages: (1) It has excellent working ability in weak light and harsh environments. (2) The target can be detected even in the presence of non-biological obstacles, with strong anti-interference ability. (3) The infrared wavelength is short, and the target image with high resolution can be obtained. (4) It

has strong camouflage target recognition ability. Consequently, enhancing the detection effect of infrared small targets in complex backgrounds can improve the perception ability of autonomous systems, so as to promote their safe operation. Consequently, enhancing the detection ability of infrared small targets has become an important research direction of autonomous system.

Hyperspectral images contain rich information, but too much information will also have a great impact on the target [1]. The imaging conditions of spectral images are rough and contain a large number of mixed pixels [2]. In contrast, infrared images have the advantages of fast imaging speed, clear imaging and more prominent target heat information. This makes infrared target detection faster and more accurate, so the application scenarios are more extensive.

Infrared targets imaging depends on the temperature difference and emissivity difference between the target itself and its surrounding environment, so the target appears in the form of highlight in the background. In infrared small-target detection, researchers have done a great deal of research, and there are many mature algorithms. However, there are three reasons for detection failure:

- (1) The imaging distance of small infrared target usually makes the pixel ratio of small target to the whole image very small.
- (2) The target radiance decreases with the increase of the action distance, which makes the target weak and the distance from the environment is low. The target is easy to be submerged by the complex background, resulting in the failure of detection.
- (3) Among other factors, the existence of interference objects similar to the target in complex imaging environment and complex background will result in a high rate of false alarms.

Therefore, infrared small-target detection still has high research value.

Nowadays, the idea of deep learning has been to detect small infrared targets and has made important contributions. The general idea of deep learning algorithm is to learn target features in a data-driven way. However, it still has the following disadvantages:

- (1) Background clutter and noise obscure small targets, resulting in their failure to be detected;
- (2) The target is very small, resulting in detection failure.

As shown in Figure 1, in these two difficult situations, the effect of the traditional algorithm is not ideal. Especially in the second difficult case, the target occupies only a few pixels. The existing algorithms have difficulty to solve this kind of problem effectively. Therefore, improving the detection ability of the algorithm in difficult situations is the key problem to be solved.

Therefore, to solve these challenges, an image enhancement algorithm combining sharpening the spatial filter and upsampling is proposed in this paper. First, the image is sharpened by spatial filter to enhance the grey mutation and the details of the image, improve the separation degree between the small target and the background and enhance the characteristics of the small target. Thus, the detection failure caused by background clutter and noise inundation is avoided. Then, the image is upsampled, and double-triple interpolation is used to increase the target pixels, so as to enlarge the image to four times its original size in equal proportion; the purpose of this is to enlarge the target. By enhancing the information on the small targets and increasing the number of pixels of the small targets, our algorithm can provide better detection in complex backgrounds. As a summary, this paper has the following contributions.

- (1) Sharpening spatial filters is proposed to enhance small targets. By increasing the contrast between the edges of the object and the surrounding image elements, the small target is emphasised at a nuanced level, and the separation of the small target from the background is increased. Compared with existing algorithms, our algorithm makes small targets clearer and easier to detect, thus solving the problem of the inaccurate detection of small targets due to background clutter and noise drowning.

- (2) Upsampling is designed to expand the target pixels, i.e., the image is scaled equally using bi-triple interpolation, allowing the enhanced small targets to be scaled up as well. Very small targets that are difficult to detect are enlarged to targets that are relatively easy to detect. In practice, the algorithm enhances the recognition of small targets and solves the problem of detection failure due to a small target.
- (3) By comparing the experiments with other algorithms on the NUAA-SIRST dataset, it is demonstrated that the algorithm proposed in this paper has better performance relative to existing algorithms in the three evaluation metrics of Pd, Fa and IoU, and it has better applications in the field of perception of autonomous systems.
- (4) Throughout the paper, we review relevant works in Section 2, introduce our proposed method in Section 3. In Section 4, the proposed detection architecture is analysed quantitatively and qualitatively, and the conclusion is given in Section 5.

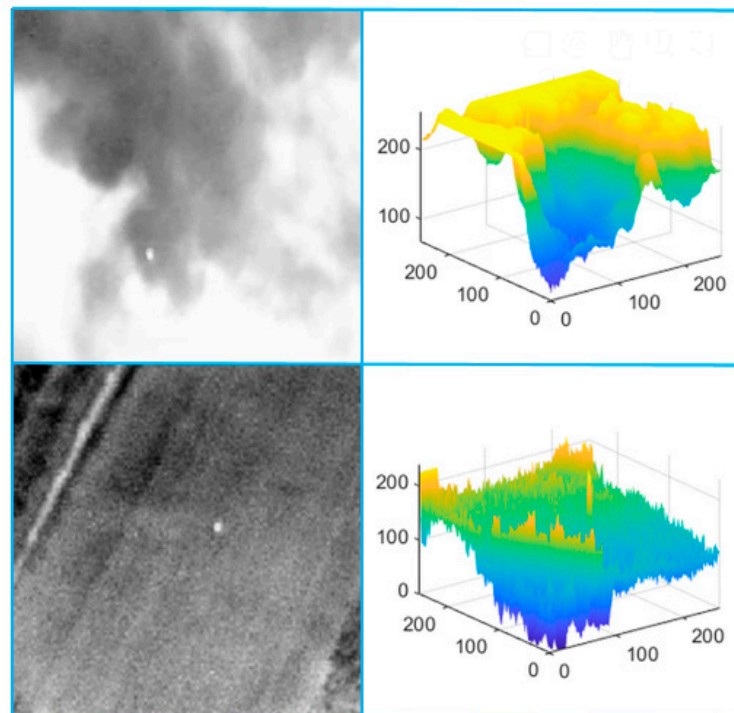


Figure 1. Background clutter and noise drowning.

2. Related Works

2.1. Infrared Small-Target Detection in Autonomous Systems

For detecting small infrared targets, a number of excellent conventional algorithms have been developed. Among them are methods based on filters [3], local contrast [4–6] and low-rank [7–9]. However, the shortcoming of these methods is that they rely too much on handcrafted features, which makes detection with handcrafted features and fixed hyperparameters difficult to work in scenarios where object size, object shape, SCR and clutter background vary significantly.

With the research on deep learning, infrared small-target detection has also started to use deep learning methods. Liu et al. [10] were the first to apply deep learning methods to the detection of small targets in infrared. They added randomly generated target points to a signal-to-noise controlled background to generate target samples for training prediction. Then, McIntosh et al. [11] improved several deep learning networks, such as Faster-RCNN [12] and Yolo-v3 [13], to improve their suitability for detecting infrared small targets.

In recent years, Park et al. have proposed a pixel-level classifier based on a convolutional neural network (CNN) for human detection in infrared closed-circuit television at night and achieved better results than traditional algorithms [14]. Hou et al. proposed a robust infrared small-target detection network (RISTDnet) based on deep learning that

constructed a feature-extraction framework combining a manual feature method and a convolutional neural network [15]. In order to focus on the basic characteristics of infrared small targets, bin Zhao et al. proposed a new detection mode based on the generation of a countermeasures network [16]. Zhang proposed a self-regularized weighted sparsity (SRWS) model to transform detection problems into optimization problems [17].

In 2021, Dai et al. designed an asymmetric contextual module (ACM) [18] optimising the downsampling scheme, attention module and feature-fusion methods. They then proposed a new network on top of their ACM. They combined discriminative networks with traditional model-driven methods to propose a model-driven deep network (ALCNet) [19], with good results.

Image super-resolution technology is also applied to image segmentation [20,21], but different from super resolution, the algorithm proposed in this paper emphasizes small targets from the subtle level, focusing on strengthening the edge of small targets, improving the separation between small targets and the background and expanding the number of target pixels that meet the DNANet prediction conditions on this basis, so as to improve the detection effect.

At this stage, deep learning was used to create a dense nested attention network (DNANet) by Li et al. [22] that was specially designed to detect small infrared targets and retained the characteristics of small targets in the deep network. Compared with the above algorithms based on deep learning, this network has achieved the most satisfactory results at this stage. These methods are all improvements on how to retain the characteristics of small targets in the deep network. However, in the case of minimal target and noise inundation, the existing algorithms cannot achieve satisfactory recognition results. In contrast, in our method in that we enhance the small target by image enhancement, so we obtain more satisfactory results in both cases.

2.2. Sharpening Spatial Filters

In the area of image sharpening, there has been much research. Asokan et al. have optimally tuned a bilateral filter [23] that reduces noise in the image and improves image quality and is good for edge preservation. Gupta et al. have designed a fractional-order digital FIR filter (ABC-FODF) [24] that is suitable for processing the signal and for various filtering operations. Peng et al. proposed a denoising sparse self-encoder-based fault-detection method for understanding input–output mapping and correlation in denoising sparse self-encoders by combining smoothed integral gradients [25]. Li et al. used a bi-directional filter to persist the edges of the fused image, a fractional difference technique for reducing noise in the fused image and low-frequency detail in smoothed areas to enhance texture detail [26]. Ghani et al. proposed a technique combining background enhancement filters and wavelet fusion for improving the contrast and visibility of underwater images [27]. The advantage is that it reduces blur and improves contrast and visibility. An image fusion method based on DWT's fusion and Roberts's operator was proposed by Paramanandham and Rajendiran. The method uses wavelet transform to obtain the initial fused image and uses Roberts's operator to extract edge information from the input image and replace the edge information in the original image with this information for the purpose of enhancing the fused image [28].

The main purpose of sharpening is to highlight the transition part of the gray scale. It compensates the contour, enhances the edge and gray jump part of the image and makes the image clear. Its main point is to enhance the part of the image that changes violently and suppress the part that changes slowly, so as to enhance the information of the image.

2.3. Upsampling

There has been considerable research on upsampling. Qi et al. proposed a method to super-resolve a single frame image. The method first constructs the optimal fractional order gradient based on image similarity and then reconstructs it using a method based on the minimum energy function to obtain a high-resolution image [29]. Malayil et al.

propose a reversible watermarking scheme based on image scaling. This method duplicates neighbouring pixels and uses them as the intensity values of the missing pixels in the newly scaled image to insert the CAPTCHA and EPR data [30]. Xue et al. introduced image-interpolation techniques for the purpose of improving the efficiency and robustness of target-detection techniques [31]. Kok et al. pointed out that among the existing algorithms, the most commonly used interpolation techniques are bilinear interpolation and bicubic interpolation [32] for the downsampling and upsampling of images, respectively. Jiang et al. achieved the desired results by extracting a series of foreground sub-images with different resolutions for the detector to reduce the computational effort and BER [33]. Davide Mazzini proposes a guided upsampling module for the efficient exploitation of high-resolution cues during upsampling [34]. Zhang et al. use grouped upsampling for extracting contextual information [35].

When the image information is added by enlarging the image, samples of the image are required. The enlarged image can be displayed on a higher-resolution device. The existing upsampling algorithms basically use interpolation. After the original image is enlarged, the value of additional pixels needs to be added in an appropriate way to expand the target pixels. This process is called interpolation.

3. Proposed Method

3.1. Revisiting DNANet Detector

The DNANet algorithm first pre-processes the input image and feeds it into a densely nested interaction module backbone to extract multilayer features. The fusion of multilayer features is then repeated at the intermediate convolutional nodes of the hopping connection, and the fused features are then output to the decoder subnet. The multilayer features are then adaptively enhanced using the channel space attention module to achieve better feature fusion. The feature pyramid fusion module connects shallow features with different types of information to deeper features, resulting in richer, more informative features and ultimately a robust feature map as shown in Equation (1), where $L_{\text{en_up}}^{iJ} \in \mathbb{R}^{C^i \times H_0 \times W_0}$ $i \in \{0, 1, \dots, I\}$ is the obtained features at all levels and G is the final obtained feature map. Finally, the feature map is fed into the eight-connected neighbourhood-clustering module, and if any two pixel points $g(m_0, n_0), g(m_1, n_1)$ have intersecting regions in their eight neighbourhoods (as in Equation (2)) and have the same value (0 or 1) (as in Equation (3)), then the spatial location of the target centre of mass is calculated and the target is predicted. The specific parameters of DNANet are shown in Table 1.

$$G = \{L_{\text{en_up}}^{0J}, L_{\text{en_up}}^{1J}, \dots, L_{\text{en_up}}^{IJ}\} \quad (1)$$

$$N_8(m_0, n_0) \cap N_8(m_1, n_1) \neq \emptyset \quad (2)$$

$$g(m_0, n_0) = g(m_1, n_1), \forall g(m_0, n_0), g(m_1, n_1) \in G \quad (3)$$

Table 1. The specific parameters of DNANet.

Stage	Conv	Max Pool	Up-Conv	Backbone	Lerning Rate
	3×3	2×2	2×2	resnet_18	0.005

However, the algorithm has certain drawbacks in some cases. When an extremely small and faint target is recognised, this makes the target pixel value $g(m_1, n_1) = 1$ and its neighbouring pixel values $g(m_0, n_0) = 0$. This makes $N_8(m_0, n_0) \cap N_8(m_1, n_1)$ not be empty, but $g(m_0, n_0) \neq g(m_1, n_1)$, i.e., the prediction condition is not satisfied, resulting in the inability to compute the spatial location of the target's centre of mass and thus predict such targets. The robustness of the algorithm in the face of such cases needs to be improved, and the following improvements are made in two aspects of this paper.

3.2. Target Feature Enhancement Based on Sharpening Spatial Filters

In Section 2, we introduced a sharpening spatial filters. We found that the idea of using sharpening spatial filters based on second-order differential-Laplace operators can solve such problems.

When DNANet predicts the input image, it needs to meet the conditions of Equations (2) and (3) to predict the target, but some targets do not meet the conditions. In order to solve this problem, the sharpening spatial filter is used to convert the values of adjacent pixels into the same values as the target pixels, making them meet Equations (2) and (3). Since in a binary map, the target pixel value $g(m_1, n_1) = 1$, we also need an algorithmic calculation of the neighbouring pixel values such that the processed $g(m_0, n_0) = 1$ to satisfy the prediction conditions so that such targets can be predicted.

In DNANet, the algorithm ultimately predicts the binary image. Therefore, to achieve the above purpose, it is necessary to increase the pixel value of the target edge so that there are more pixels with a pixel value of 1 around the target after conversion to a binary image. When the value of a pixel point is greater than ε , it is set to 1, and when it is less than ε , it is set to 0. Because the pixel values of the target edges fall in a gradient, there is a grey scale step. This results in some pixels at the edge of the target having a value less than but very close to ε , $f(m_0, n_0) < \varepsilon$. This makes it impossible for the points adjacent to the target pixel to satisfy the constraint. Since property two of second-order differentiation is known, the differentiation is not zero at the start of the grey scale step or slope, i.e., $\nabla^2 f(m_0, n_0) > 0$ using the second-order differential-Laplace operator, which is calculated as Equation (4):

$$F(x, y) = f(x, y) + c [\nabla^2 f(x, y)] \quad (4)$$

where $f(x, y)$ and $F(x, y)$ are the input image and the sharpened image, respectively. For the purpose of making the enhanced image retain the background characteristics as well as the sharpening effect, the central factor $c = 1$ is taken in this paper.

Taking the point (m_0, n_0) into Equation (4) gives $F(m_0, n_0) = f(m_0, n_0) + [\nabla^2 f(m_0, n_0)] > \varepsilon$. At this point, $N_8(m_0, n_0) \cap N_8(m_1, n_1)$ is not empty, and after transforming into a binary image, $g(m_0, n_0) = g(m_1, n_1) = 1$, which satisfies the constraint, thus solving the above problem and achieving the recognition of such targets.

The effect of the edges of the small target after sharpening and filtering enhancement is shown in Figure 2.

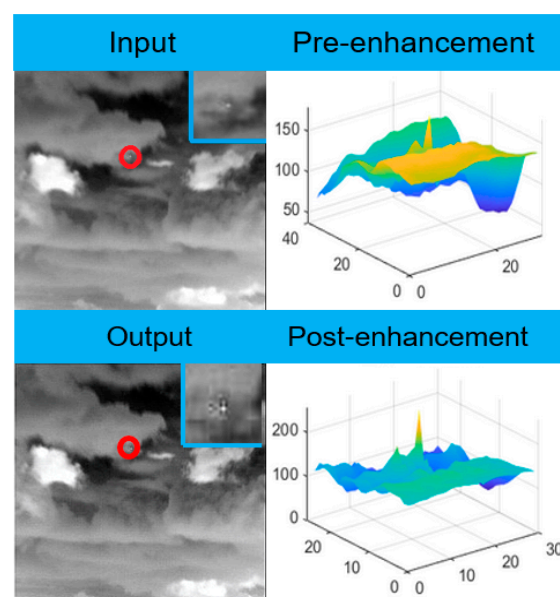


Figure 2. 3D visualisation results for the target before and after enhancement.

Using Figure 2 as an example, the signal-to-noise ratio (SCR) of this small target is calculated. A small target whose SCR is high is easier to detect. As can be seen from Table 2, after the enhancement of our algorithm m , the SCR of the small target is enhanced from 0.074 to 0.134, an improvement of 0.06.

Table 2. Signal-to-noise ratios (SCRs) before and after small-target enhancement.

	Pre-Enhancement	After Enhancement
SCR	0.074	0.134

3.3. Target Pixel Point Expansion Based on Upsampling

Only the image has been sharpened and filtered. Although it has brought some improvements in detection, there is still some room for improvement. Therefore, we need to add pixels in the image on the basis of image enhancement. The purpose of this step is to make more pixels meet the constraint condition of $g(m_0, n_0) = g(m_1, n_1) = 1$ on the basis of $N_8(m_0, n_0) \cap N_8(m_1, n_1) \neq \emptyset$, so as to further optimize the detection effect. Therefore, we use upsampling to achieve this purpose.

For the purpose of making the expanded pixel points closer to the original image, i.e., to better preserve the small targets after enhancement, the algorithm in this paper uses bicubic interpolation to increase the number of pixels in the image. This approach results in an enlargement closer to the high-resolution image, which is calculated as follows.

First construct the bicubic function:

$$W(x) = \begin{cases} |x|^3 - 2|x|^2 + 1 & |x| \leq 1 \\ -|x|^3 + 5|x|^2 + 4 & 1 < |x| < 2 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

As shown in Figure 3, x is the distance from each pixel point to point p . The weights $W(x)$ corresponding to the 16 pixels around point p are obtained by finding the parameter x . When the weights of the 16 pixel points around the target pixel point p are obtained, the value of the enlarged image (x, y) is equal to the weighted superposition of the 16 pixel points, which is calculated as follows:

$$B(X, Y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} \times W(i) \times W(j). \tag{6}$$

where a_{ij} is the location of 16 pixel points. By expanding the target pixel points, because the pixel values obtained from the bi-triple interpolation calculation are closest to the pixel values of the original image, more pixel points meet the constraints, thus circumventing the problem of not detecting the target due to its being too small and improving the detection effect.

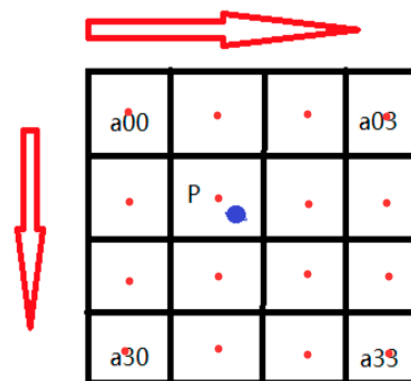


Figure 3. Pixel value calculation.

As shown in Figure 4, after the upsampling process, the number of small target pixels is increased, and more pixel points satisfy the constraints.

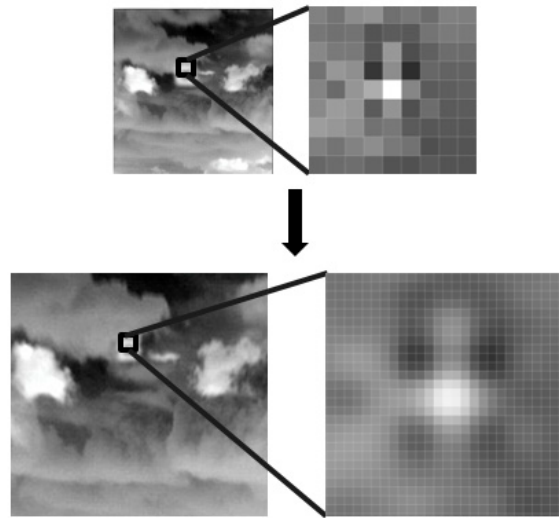


Figure 4. Expanded target pixels.

3.4. Overall Process

We propose in this paper an image enhancement-based algorithm to detect infrared small targets. This is because some small targets themselves do not have a prominent grey-scale jump in the area connected to the background, or the number of pixels in the small target is very small compared with the number of pixels in the whole image. In both cases, the number of pixels that satisfy the prediction criteria for the small target itself is insufficient. Existing algorithms are not effective in detecting these two cases. Therefore, this paper starts with improving the detection head of the algorithm by enhancing the small targets to improve the detection effect. The first step is to use the idea of sharpening spatial filtering to improve the pixel value of the target. On this basis, the enhanced image is upsampled to increase the number of target pixels. This avoids the problem of detection failure due to background clutter, noise drowning and small targets. The overall flow chart of the method is shown in Figure 5.

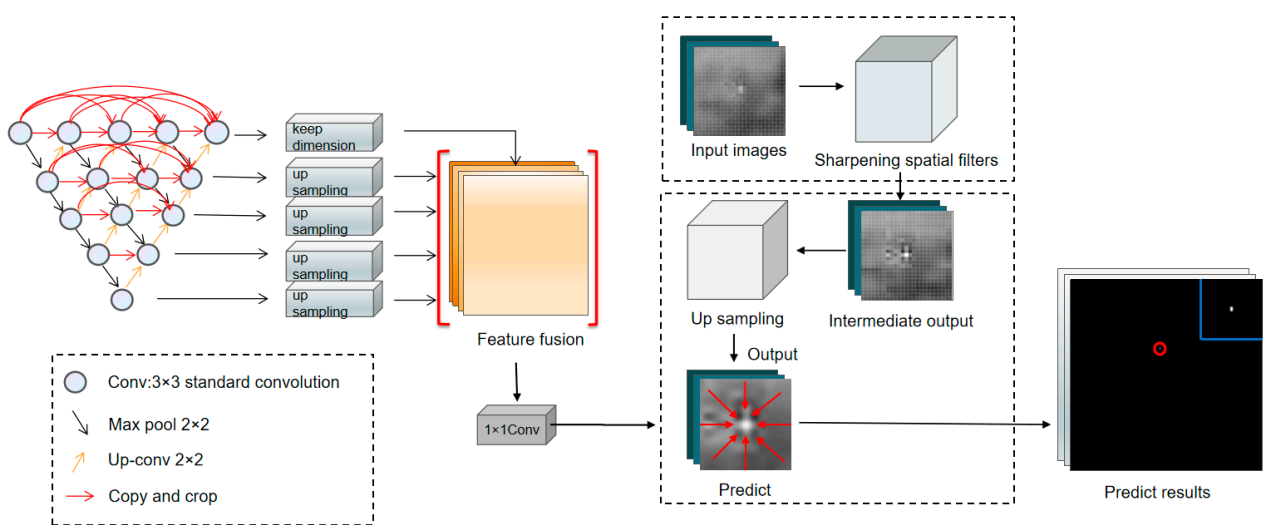


Figure 5. Overall flow chart.

4. Results of the Numerical Experiments

4.1. Introduction to the Dataset

In this section, we use the NUAA-SIRST dataset [18] as a benchmark to compare the results of our proposed algorithm with those of existing algorithms. In total, there are 427 infrared images in the dataset, of which there are 480 instances, and the dataset is divided into three parts: a training set, a validation set and a test set in the ratio 5:2:3. Figure 6 shows a preview of the dataset in which the target has low contrast with the background and is easily swamped by a complex background with heavy clutter. Here, even with the naked eye, such targets are hard to discern from the background. Most images in the dataset contain only one target, and only a few images contain multiple targets. The dataset is classified according to the target number, and Figure 7 is obtained. Approximately 64% of these targets are only 0.05% around of the whole image. The vast majority of the targets are less than 0.1% and are classified by the percentage of targets to obtain Figure 8. Only 35% of the targets are the brightest in the image, so most of the targets in this dataset are very small and faint.

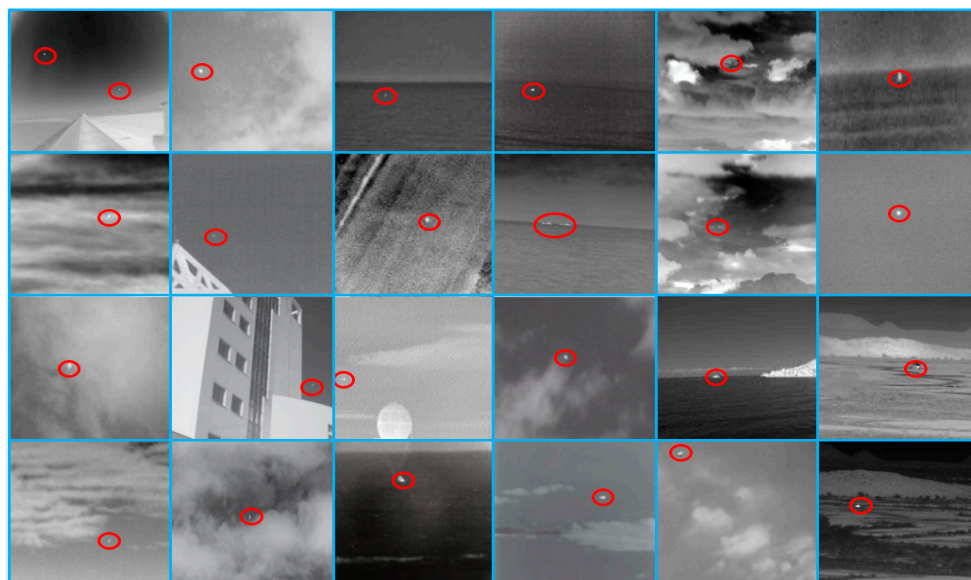


Figure 6. NUAA-SIRST dataset, where the red circles are the targets.

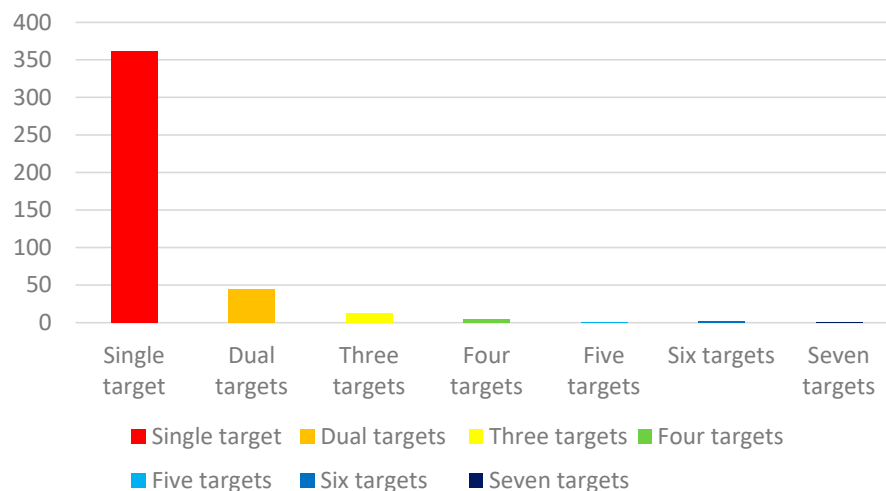


Figure 7. Breakdown by number of targets.

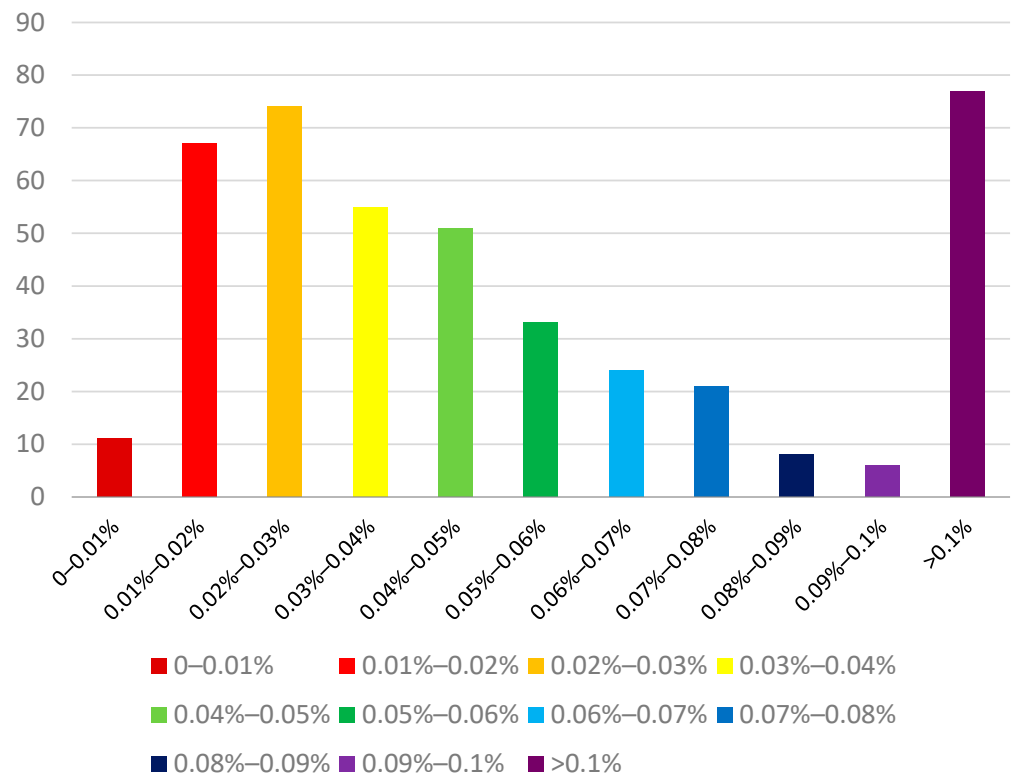


Figure 8. Breakdown by percentage of targets.

4.2. Assessment Indicators

CNN-based algorithms [18,19,36] mainly use pixel-level evaluation metrics such as IoU and accuracy. These metrics are mainly focused on target shape evaluation. Algorithms based on local contrast metrics mainly use Pd and Fa as the evaluation metrics to assess the detection results. In the DNANet [22] algorithm, the three metrics IoU, Pd and Fa are combined as evaluation metrics. This is because Pd and Fa can evaluate the detection results more intuitively, and IoU is more indicative of the difference between the size and coordinates of the detected target and the ground truth. Therefore, IoU, Pd, and Fa are used as evaluation indicators in this paper.

- (1) IoU (Intersection of Union) is a standard to measure the detection accuracy, which can evaluate the shape detection ability of the algorithm. The result can be obtained by calculating the overlap between the predicted target and the ground truth value divided by the union of the two regions, where Area of Overlap represents the overlap and Area of Union represents the union part:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of union}} \quad (7)$$

- (2) Detection rate Pd: Pd measures the accuracy of target detection by comparing the detected results with ground truth to Pd, where the number of targets detected is N_p and the number of ground truths is N_r :

$$\text{Pd} = \frac{N_p}{N_r} \quad (8)$$

- (3) False alarm rate Fa:FA is used to evaluate the degree of misjudgement. The result is obtained by calculating the ratio of mispredicted pixels to all pixels of the image, where PF is the misjudged pixels and PA is all pixels in the image:

$$Fa = \frac{Pf}{Pa} \tag{9}$$

- (4) Mean Intersection over Union:mIoU is an index used to measure the accuracy of image segmentation. The higher the mIoU, the better the performance. Intersection is the number of pixels in the intersection area, and combine is the number of pixels in the union area.

$$mIoU = \frac{1}{K} \sum_{i=1}^K \frac{\text{Intersection}}{\text{Combine}} \tag{10}$$

4.3. Quantitative Analysis

We compared the algorithms presented in this paper with the current state-of-the-art algorithms, including TLLCM, tri-layer local contrast measure [5]; WSLCM, weighted strengthened local contrast measure [6]; RIPT, teweighted infrared patch-tensor model [7]; NRAM, non-convex rank approximation minimization joint [8]; PSTNN, partial sum of the tensor nuclear norm [9]; ALCNet, attentional local contrast networks [19]; DNANet, dense nested attention network [22]; MDvsFAcGAN, missed detection vs. false alarm; conditional generative adversarial network [36]; MSLSTIPT, multiple subspace learning and spatial-temporal patch-tensor model [37]; and MPANet, multi-patch attention network [38]. These algorithms and our algorithm were tested on the NUAA–SIRST dataset, and the obtained quantitative and qualitative results are compared and analysed. The quantitative analysis focuses on the changes in the three metrics of IoU, Pd and Fa between the algorithms proposed in this paper and those mentioned above, so as to see more intuitively how much the detection effectiveness of the algorithms has improved, as shown in Table 3 and Figure 9.

Table 3. IoU, Pd and Fa obtained by the different algorithms on the NUAA–SIRST dataset.

Metric	WSLCM	TLLCM	NRAM	RIPT	PSTNN	MSLSTIPT	MDvsFA-cGAN	ALCNet	DNANet	Proposed
IoU ($\times 10^{-2}$)	1.158	1.029	12.16	11.05	22.40	10.30	60.30	73.33	75.46	75.55
Pd ($\times 10^{-2}$)	77.95	79.09	74.52	79.08	77.95	82.13	89.35	95.57	96.95	98.48
Fa ($\times 10^{-6}$)	5446	5899	13.85	22.01	29.11	1131	56.35	30.47	13.23	10.09

Higher IoU and Pd indicate higher performance. Smaller Fa indicates higher performance. The best values are marked in red, the second-best values in blue and the third-best values in yellow.

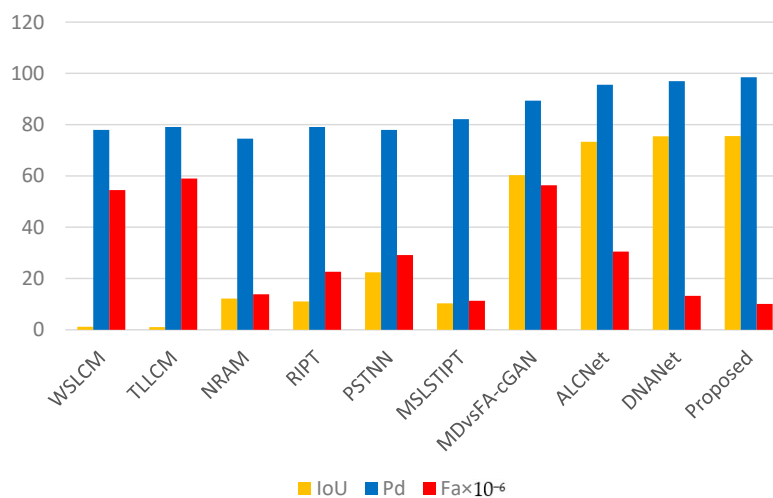


Figure 9. Comparison of our algorithm with other algorithms.

Table 3 shows an improvement in the overall rate. Compared with current advanced methods, the algorithm proposed in this paper makes a certain improvement in IoU over the algorithm DNANet, indicating that our algorithm has a strong ability to describe the target contour on the basis of the detected target. The stronger the ability to describe the target contour, the easier it is to determine the type of small targets. Thus, it improves the perception ability of the autonomous system and provides more judgment basis for the decision-making of the autonomous system.

The 1.14% improvement in detection rate for Pd over DNANet proves that the improvements made by our algorithm over the original algorithm are effective. The improvement in the most important metric, Pd, demonstrates that we improved the detection capability of the algorithm by augmenting it with tiny and faint targets while not misclassifying small target analogues and leading to higher false alarm rates. The improvement in the detection rate Pd greatly ensures the safe operation of the autonomous system.

The false alarm rate Fa is reduced by 2.4×10^{-6} compared with DNANet, and the input image is processed by our algorithm, and for some small target analogues that are easily misjudged, the characteristics similar to those of the small target are weakened, thereby reducing the misjudgement of the small target.

Combining these three indicators, it is verified that our algorithm can not only improve detection while ensuring a low false alarm rate after enhancing the input image, it also has a strong ability to describe the target contour. The application of our algorithm to an autonomous system can effectively improve the perception capability of the self-help system.

4.4. Quantitative Analysis

Qualitative analysis mainly analyses whether the image enhancement module proposed in this paper solves the problems of DNANet and verifies our theory in Section 3. We more intuitively show whether our algorithm will detect small targets that are difficult to detect and whether the recognition effect in difficult situations will be better than DNANet.

4.4.1. Enhanced Target Characteristics

A solution to the problem of resolving detection failure caused by small targets submerged by background noise or noise is proposed in Section 1. We first perform sharpening filtering on the input image to enhance the edge information of the image. As shown in Figure 10, the feature maps before and after enhancement show that the peaks are improved after enhancement and the surrounding background information is suppressed, proving that the features of the small target itself are enhanced and thus easier to detect. The output also shows that the brightness of the target has improved and the edges are clearer than in the original image. Figure 11 also shows that after sharpening, the brightness of the small target is increased and the edge information is enhanced, making it easier to detect than the original image. This confirms the validity of the method mentioned in Section 3 and thus solves the above problem.

4.4.2. Expanded Target Pixels

The problem with the target proposed in the first section is that it has a very small proportion compared with the background, which leads to the failure of detection. As shown in Figure 12, we perform pixel expansion on the small target by upsampling based on the enhancement of the small target, and the target becomes rich in content. As can be seen in Figure 13, the recognition of the expanded target pixels gives a result that is closer to the true value of the silhouette and improves the accuracy of the recognition compared with the original image. This shows that our proposed method can solve this kind of problem ideally.

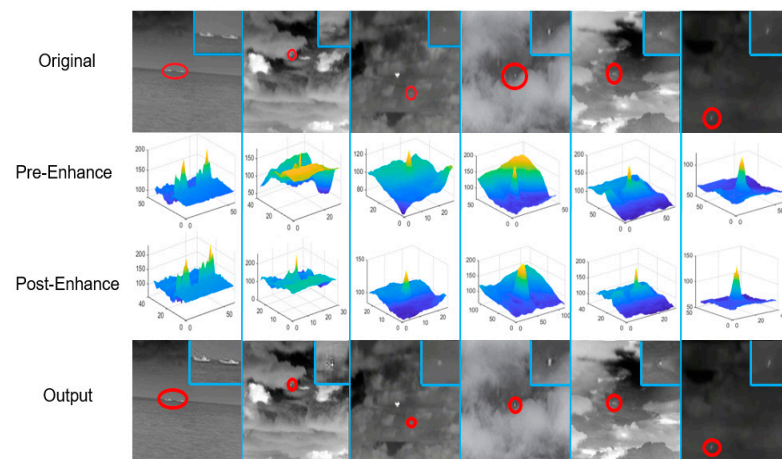


Figure 10. Enhancing small targets.

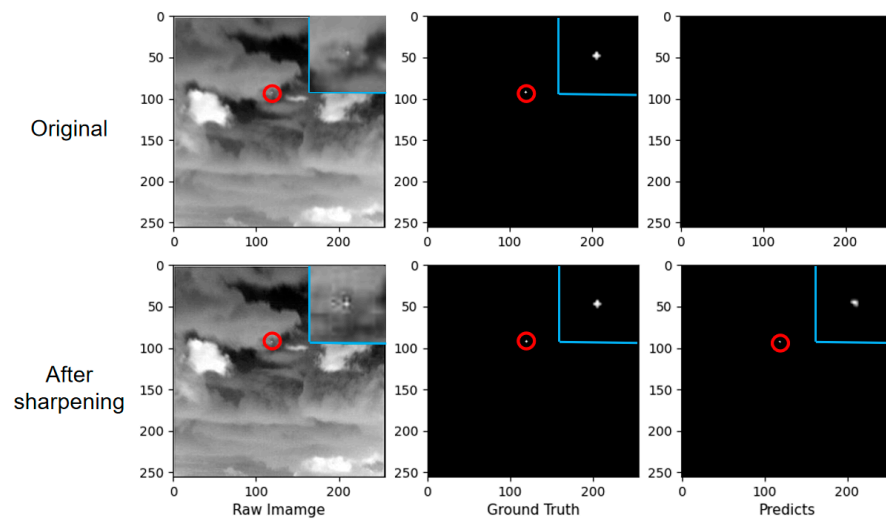


Figure 11. Detection results before and after sharpening.

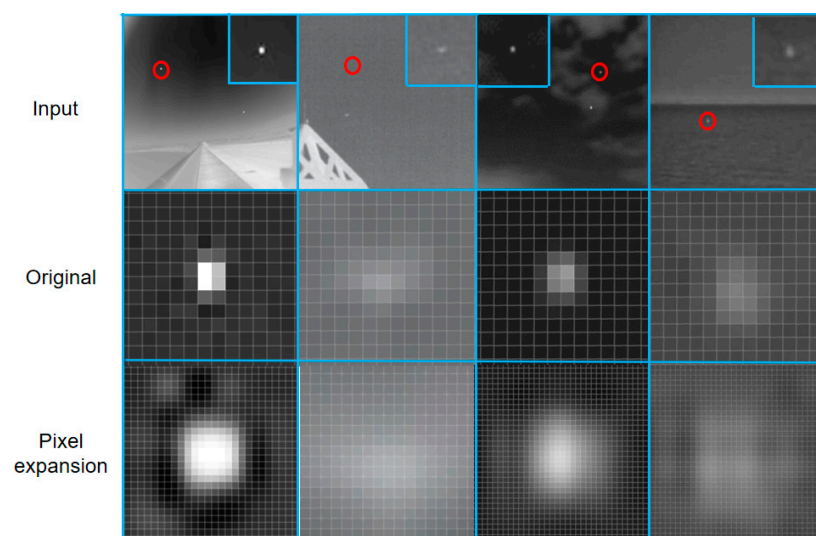


Figure 12. Experiment with expanded target pixels.

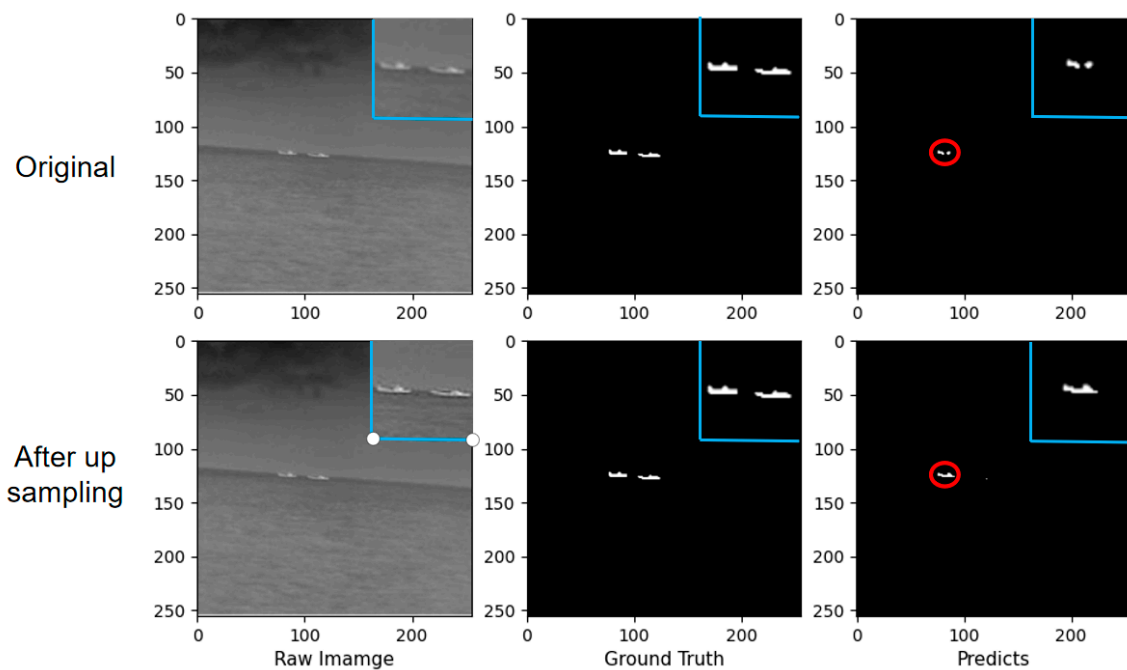


Figure 13. Detection results before and after upsampling.

4.4.3. Comparison of Test Results

As shown in Figure 14, we compare the detection results after target feature enhancement and pixel expansion with those for the DNANet algorithm. It can be seen that there is an overall improvement in detection after the processing of our algorithm.

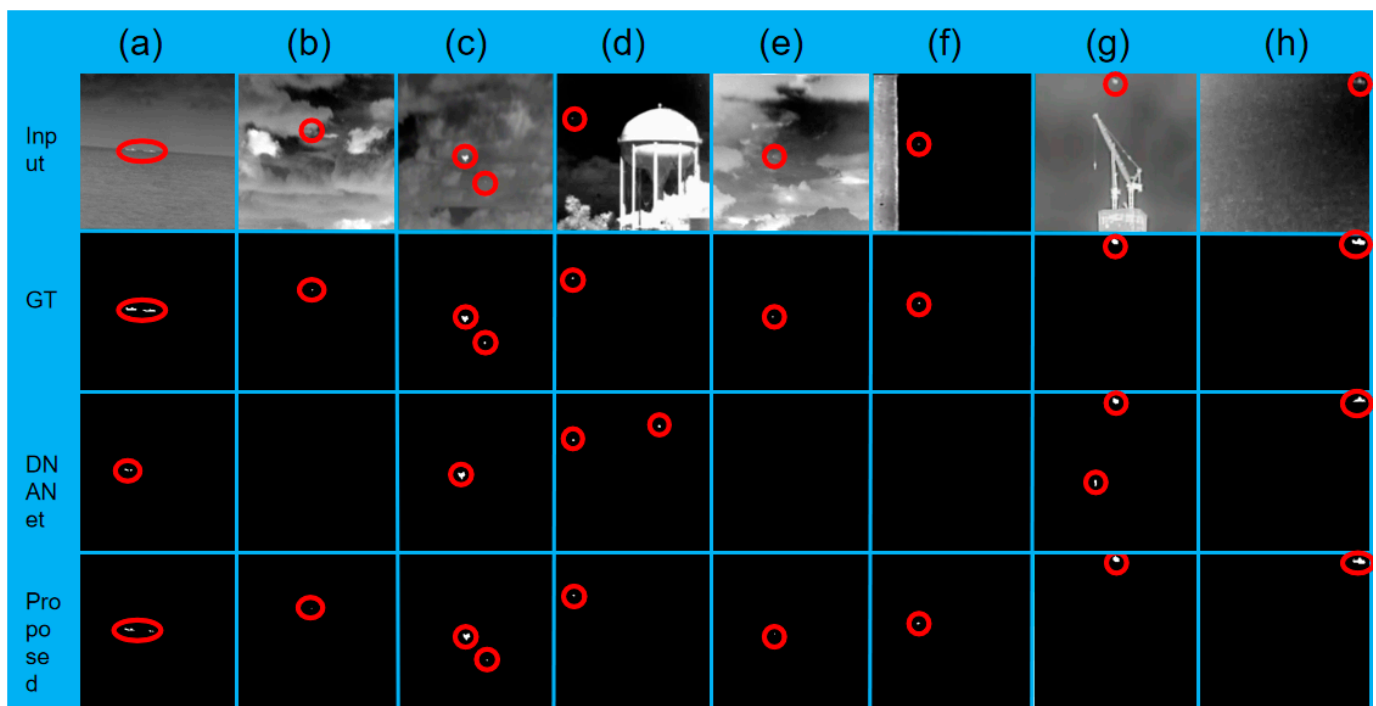


Figure 14. Comparison of detection results with DNANet, in which the subfigures (a–h) are some experimental samples in recognition with red circles containing targets, the 1st line shows original pictures, 2nd line shows the ground truth of the recognition results, 3rd line shows the recognition results of DNANet, and the last line shows the recognition result of the proposed method in this paper.

The results show that in row (a), DNANet has a problem of inadequate recognition of this target, recognizing only part of the target due to the fact that some pixel values at the edges of the target are slightly below the banalization threshold ϵ . The number of pixels that meet the constraints is small. After processing by our algorithm, the pixel values of the target edges were increased and the pixel points were expanded, resulting in some improvement in the detection results.

In row (b), (e), (f), which is a typical example of a very small and dim target, the human eye also has difficulty finding this target from the background, and DNANet does not recognize such targets well enough to miss them, resulting in a decrease in the detection rate P_d . However, after processing by our algorithm, it is possible to enhance such very small targets, i.e., point targets, which are difficult to detect. The pixel values of the target edges are first increased so that the small targets stand out more compared with the background, and then the pixels of the small targets are expanded so that they are easier to detect, enabling the recognition of such targets and resulting in an increased detection rate.

From line (c), it can be seen that there are two small targets in the original image, one of which is larger and brighter, and the other is very small and faint. It is clear that DNANet can achieve accurate recognition for the larger and brighter target, but it has difficulty detecting the other target. In (c), a comparison shows that DNANet does have this problem and that our algorithm can solve it effectively.

In row (d), (g), it is the other small target analogues in the graph that DNANet misidentifies, resulting in a higher false alarm rate. This is because the analogue has similar features to the small target in the original image, but after our algorithm has increased its edge pixel value and expanded the pixels, it weakens its similar features to the small target, allowing the algorithm to detect it without false positives, thus reducing the false alarm rate.

To sum up, our proposed algorithm can effectively solve the two problems of (1) background clutter and noise swamping small targets, leading to detection failure and (2) targets accounting for too little compared with the background, leading to detection failure, thus improving the reliability and robustness of the algorithm in complex environments and making an important contribution to the improvement of the sensing capability of autonomous systems.

Table 4 corresponds to the columns in Figure 13. We can see that compared with DNANet, the proposed algorithm achieves more satisfactory results for mIoU, which proves that the proposed algorithm has better segmentation performance.

Table 4. Comparison of mIoU for the proposed method with that of DNANet.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
DNANet	0.21	0	0.74	0.28	0	0	0.80	0.65
Proposed	0.43	0.4	0.84	0.88	0.37	0.58	0.90	0.83

4.5. Ablation Experiments

In this paper, the method of sharpening first and then upsampling is adopted. The purpose is to first use sharpening spatial filtering to improve the pixel value of the target edge and to then expand the pixel points of the enhanced image, so as to increase the number of pixels that meet the constraint conditions of $g(m_0, n_0) = g(m_1, n_1) = 1$ and achieve the recognition of small and dim targets and improve the detection effect. To test our idea, we performed the following ablation studies in this section.

We compared the algorithm in this paper with the three methods using only the sharpening spatial filter, upsampling only and upsampling first and then using the sharpening spatial filter; we further explain why the sharpening first and then upsampling approach is used:

- (1) Use of sharpening spatial filters only: Only the mutation information, details and edge information of the image are enhanced, without upsampling.

- (2) Upsampling only: cubic interpolation on the matrix of the input image using bicubic filtering only to increase the target pixel.
- (3) Upsampling and then using the sharpening spatial filter: first upsampling the image to increase the target pixels and then using the sharpening spatial filter.

According to the ablation study in Table 5, the detection rate Pd is reduced by 0.77%, the IoU by 0.54% and the false alarm rate Fa by 1.07×10^{-6} if only the sharpening is performed using the sharpening space filter. This is because if the image is only sharpened, although the pixel value of the edge of the small target will be increased, when the target proportion is extremely small, the number of pixels that meet the constraints is still not enough; therefore, this method has certain limitations.

Table 5. Ablation studies.

	IoU ($\times 10^{-2}$)	Pd ($\times 10^{-2}$)	Fa ($\times 10^{-6}$)
Sharpening only	75.01	97.71	11.97
Up-sampling only	75.22	96.95	12.12
Up-sample then sharpen	75.42	96.95	12.19
Proposed	75.55	98.48	10.90

If the image is only upsampled, the detection rate Pd decreases by 1.53%, the IoU decreases by 0.33%, and the false alarm rate Fa increases by 1.22×10^{-6} . This is because upsampling the image only expands the target pixels compared with the unenhanced target, but there is no increase in the number of pixels for which the small target satisfies the constraints, i.e., the small target itself is not enhanced. The detection rate Pd is not improved in this way, and instead, targets are detected that were not easily misidentified before.

Therefore, we combined the two approaches, upsampling the image first and then sharpening it. From the results, it can be seen that there is no certain improvement in the detection rate Pd or false alarm rate Fa compared with upsampling only, except for a small improvement of 0.20% in IoU. However, it can be seen from sharpening only that the use of the sharpening spatial filter enhances the features of small targets to improve detection, and the reason there is no improvement here is that when the input image is upsampled and then sharpened, the original pixel values that do not satisfy the constraints are expanded. By sharpening the image afterwards, only the pixel values at the edges of the target are enhanced, so the pixels that satisfy the constraint are not expanded, which makes the result unsatisfactory.

To sum up, we adopt the method of upsampling with sharpening spatial filter first. As can be seen in Figures 15 and 16, this approach we took achieved optimal results. By sharpening the input image first, the pixel values of the target edges are increased, thus enhancing the features of small targets. After that, the upsampling operation will further expand the pixels on the basis of the enhanced small target, that is, increase the number of pixels meeting the constraints, so as to fully improve the number of pixels meeting the constraints and improve the detection ability. These two methods, when combined with the right thinking, achieve the desired results.

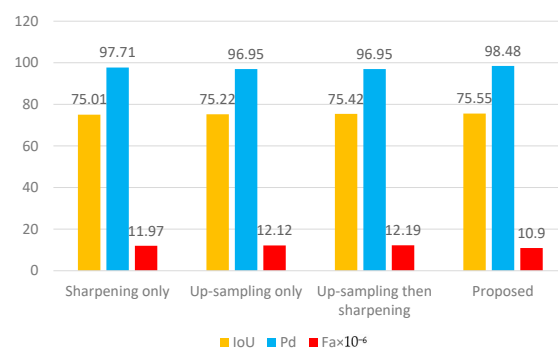


Figure 15. Ablation study.

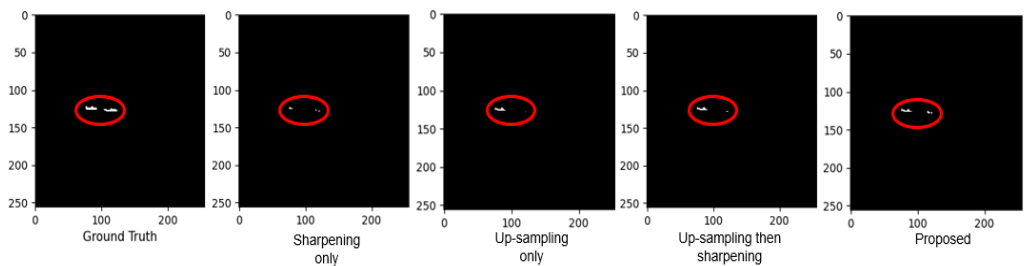


Figure 16. Comparison of ablation study results, where targets are shown in the red circle.

5. Conclusions

In this paper, we propose an algorithm for infrared small-target detection based on small-target enhancement. Unlike conventional and deep learning-based algorithms, we improve the algorithm from the perspective of enhancing small targets in the input image. We achieve this by combining both sharpening spatial filters and upsampling to improve detection. The sharpening spatial filter enhances small targets at a subtle level, making them more distinctive. The upsampling process amplifies the enhanced small targets, making difficult-to-detect point targets relatively easy to detect. The proposed algorithm effectively solves the problem that small objects are difficult to detect due to their small proportion or dimness. We compare with existing methods on public datasets and conduct extensive ablation studies. The results show that our method outperforms existing methods.

Author Contributions: Conceptualization, S.L. and M.W.; methodology, S.L. and M.W.; software, P.C.; validation, P.C. and S.L.; formal analysis, S.L. and M.W.; investigation, S.L., P.C. and M.W.; resources, S.L. and M.W.; data curation, P.C. and S.L.; writing—original draft preparation, S.L., P.C. and M.W.; writing—review and editing, S.L., P.C. and M.W.; visualization, S.L., P.C. and M.W.; supervision, S.L. and M.W.; project administration, S.L. and M.W.; funding acquisition, S.L. and M.W. All authors have read and agreed to the published version of the manuscript.

Funding: The work is sponsored by the Natural Science Foundation of Hunan Province with No. 2020JJ4434 and 2020JJ5368; Key Scientific Research Projects of Department of Education of Hunan Province with No. 19A312; Key Research Project on Degree and Graduate Education Reform of Hunan Province with No. 2020JGZD025; the National Social Science Foundation of China with No. AEA200013; and the Industry–Academic Cooperation Foundation of the Ministry of Education of China with No. HKEDU-CK-20200413-129. The authors would like to acknowledge contribution to this research from the Rector of the Silesian University of Technology, Gliwice, Poland, under pro-quality grant No. 09/010/RGJ22/0068.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Shang, X.; Song, M.; Wang, Y.; Yu, C.; Yu, H.; Li, F.; Chang, C.I. Target-constrained interference-minimized band selection for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 6044–6064. [[CrossRef](#)]
- Wang, P.; Wang, L.; Leung, H.; Zhang, G. Super-resolution mapping based on spatial–spectral correlation for spectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 2256–2268. [[CrossRef](#)]
- Rivest, J.F.; Fortin, R. Detection of dim targets in digital infrared imagery by morphological image processing. *Opt. Eng.* **1996**, *35*, 1886–1893. [[CrossRef](#)]
- Han, J.; Ma, Y.; Zhou, B.; Fan, F.; Liang, K.; Fang, Y. A robust infrared small target detection algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 2168–2172.
- Han, J.; Moradi, S.; Faramarzi, I.; Liu, C.; Zhang, H.; Zhao, Q. A local contrast method for infrared small-target detection utilizing a tri-layer window. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1822–1826. [[CrossRef](#)]
- Han, J.; Moradi, S.; Faramarzi, I.; Zhang, H.; Zhao, Q.; Zhang, X.; Li, N. Infrared small target detection based on the weighted strengthened local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1670–1674. [[CrossRef](#)]
- Dai, Y.; Wu, Y. Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3752–3767. [[CrossRef](#)]

8. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared small target detection via non-convex rank approximation minimization joint $l_2, 1$ norm. *Remote Sens.* **2018**, *10*, 1821. [[CrossRef](#)]
9. Zhang, L.; Peng, Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sens.* **2019**, *11*, 382. [[CrossRef](#)]
10. Liu, M.; Du, H.Y.; Zhao, Y.J.; Dong, L.Q.; Hui, M.; Wang, S.X. Image small target detection based on deep learning with SNR controlled sample generation. *Curr. Trends Comput. Sci. Mech. Autom.* **2017**, *1*, 211–220.
11. McIntosh, B.; Venkataramanan, S.; Mahalanobis, A. Infrared arget detection in cluttered environments by maximization of a target to clutter ratio (tcr) metric using a convolutional neural network. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *57*, 485–496. [[CrossRef](#)]
12. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
13. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
14. Park, J.; Chen, J.; Cho, Y.K.; Kang, D.Y.; Son, B.J. CNN-based person detection using infrared images for night-time intrusion warning systems. *Sensors* **2019**, *20*, 34. [[CrossRef](#)]
15. Hou, Q.; Wang, Z.; Tan, F.; Zhao, Y.; Zheng, H.; Zhang, W. RISTDnet: Robust infrared small target detection network. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [[CrossRef](#)]
16. Zhao, B.; Wang, C.; Fu, Q.; Han, Z. A novel pattern for infrared small target detection with generative adversarial network. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4481–4492. [[CrossRef](#)]
17. Zhang, T.; Peng, Z.; Wu, H.; He, Y.; Li, C.; Yang, C. Infrared small target detection via self-regularized weighted sparse model. *Neurocomputing* **2021**, *420*, 124–148. [[CrossRef](#)]
18. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Asymmetric contextual modulation for infrared small target detection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 950–959.
19. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Attentional local contrast networks for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9813–9824. [[CrossRef](#)]
20. Wang, H.; Lin, L.; Hu, H.; Chen, Q.; Li, Y.; Iwamoto, Y.; Han, X.H.; Chen, Y.W.; Tong, R. Patch-Free 3D Medical Image Segmentation Driven by Super-Resolution Technique and Self-Supervised Guidance. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2021; pp. 131–141.
21. Hakim, L.; Zheng, H.; Kurita, T. Improvement for Single Image Super-resolution and Image Segmentation by Graph Laplacian Regularizer based on Differences of Neighboring Pixels. *Int. J. Intell. Eng. Syst.* **2021**, *15*, 95–105.
22. Li, B.; Xiao, C.; Wang, L.; Wang, Y.; Lin, Z.; Li, M.; An, W.; Guo, Y. Dense nested attention network for infrared small target detection. *arXiv* **2021**, arXiv:2106.00487.
23. Asokan, A.; Anitha, J. Adaptive Cuckoo Search based optimal bilateral filtering for denoising of satellite images. *ISA Trans.* **2020**, *100*, 308–321. [[CrossRef](#)] [[PubMed](#)]
24. Gupta, A.; Kumar, S. Design of Atangana-Baleanu-Caputo fractional-order digital filter. *ISA Trans.* **2021**, *112*, 74–88. [[CrossRef](#)] [[PubMed](#)]
25. Peng, P.; Zhang, Y.; Wang, H.; Zhang, H. Towards robust and understandable fault detection and diagnosis using denoising sparse autoencoder and smooth integrated gradients. *ISA Trans.* **2021**, *125*, 371–383. [[CrossRef](#)]
26. Li, H.; Yu, Z.; Mao, C. Fractional differential and variational method for image fusion and super-resolution. *Neurocomputing* **2016**, *171*, 138–148. [[CrossRef](#)]
27. Ghani, A.S.A.; Nasir, A.F.A.; Tarmizi, W.F.W. Integration of enhanced background filtering and wavelet fusion for high visibility and detection rate of deep sea underwater image of underwater vehicle. In Proceedings of the 2017 5th International Conference on Information and Communication Technology (ICOICT), Melaka, Malaysia, 17–19 May 2017; pp. 1–6.
28. Paramanandham, N.; Rajendiran, K. Multi sensor image fusion for surveillance applications using hybrid image fusion algorithm. *Multimed. Tools Appl.* **2018**, *77*, 12405–12436. [[CrossRef](#)]
29. Yang, Q.; Zhang, Y.; Zhao, T.; Chen, Y. Single image super-resolution using self-optimizing mask via fractional-order gradient interpolation and reconstruction. *ISA Trans.* **2018**, *82*, 163–171. [[CrossRef](#)] [[PubMed](#)]
30. Malayil, M.V.; Vedhanayagam, M. A novel image scaling based reversible watermarking scheme for secure medical image transmission. *ISA Trans.* **2021**, *108*, 269–281. [[CrossRef](#)] [[PubMed](#)]
31. Xue, K.; Liu, Y.; Ogunmakin, G.; Chen, J.; Zhang, J. Panoramic Gaussian Mixture Model and large-scale range background subtraction method for PTZ camera-based surveillance systems. *Mach. Vis. Appl.* **2013**, *24*, 477–492. [[CrossRef](#)]
32. Kok, C.W.; Tam, W.S. *Digital Image Interpolation in Matlab*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
33. Jiang, Z.; Huynh, D.Q.; Moran, W.; Challa, S. Combining background subtraction and temporal persistency in pedestrian detection from static videos. In Proceedings of the 2013 IEEE International Conference on Image Processing, Melbourne, VIC, Australia, 15–18 September 2013; pp. 4141–4145.
34. Mazzini, D. Guided upsampling network for real-time semantic segmentation. *arXiv* **2018**, arXiv:1807.07466.
35. Zhang, P.; Liu, W.; Zeng, Y.; Lei, Y.; Lu, H. Looking for the detail and context devils: High-resolution salient object detection. *IEEE Trans. Image Processing* **2021**, *30*, 3204–3216. [[CrossRef](#)]
36. Wang, H.; Zhou, L.; Wang, L. Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 8509–8518.

-
37. Sun, Y.; Yang, J.; An, W. Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 3737–3752. [[CrossRef](#)]
 38. Wang, A.; Li, W.; Wu, X.; Huang, Z.; Tao, R. MPANet: Multi-Patch Attention For Infrared Small Target object Detection. *arXiv* **2022**, arXiv:2206.02120.