remote sensing

MDPI

Article

CHFNet: Curvature Half-Level Fusion Network for Single-Frame Infrared Small Target Detection

Mingjin Zhang, Bate Li, Tianyu Wang, Haichen Bai, Ke Yue and Yunsong Li *

State Key Laboratory of Integrated Services Networks, School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
* Correspondence: ysli@mail.xidian.edu.cn

Abstract: Single-frame infrared small target detection (IRSTD) aims to extract targets from background clutter and distinguish them from noise. In recent years, semantic segmentation deep learning methods such as CNNs have made many breakthroughs in the field of IRSTD. However, there are limitations to this method; for example, the targets tend to be too dim, and heavy background clutter exists. To further improve the accuracy of IRSTD, we propose a novel curvature half-level fusion network (CHFNet) for IRSTD. First, we developed a half-level fusion (HLF) block as a new cross-layer feature fusion module. With the HLF block, the network excavates the half-level features between two levels of features, thus ensuring that each feature of the levels has minimal distortion. Given that even dim targets have certain curvature features, we calculated the weighted mean curvature of the image to obtain the attention of the boundary, then fused it with the features of each level to detect the edges of targets. In comparison, the prediction results of the proposed CHFNet on the NUAA-SIRST dataset were more complete and better preserved edge targets.

Keywords: infrared small target detection; semantic segmentation; weighted mean curvature; half-level fusion

check for updates

Citation: Zhang, M.; Li, B.; Wang, T.; Bai, H.; Yue, K.; Li, Y. CHFNet: Curvature Half-Level Fusion Network for Single-Frame Infrared Small Target Detection. *Remote Sens.* **2023**, *15*, 1573. https://doi.org/10.3390/rs15061573

Academic Editor: Gemine Vivone

Received: 19 January 2023
Revised: 22 February 2023
Accepted: 27 February 2023
Published: 13 March 2023

Copyright: © 2023 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

Single-frame infrared small target detection (IRSTD) is widely used in applications such as marine rescue, forest fire patrol, natural disaster rescue, etc. There exist some traditional methods for IRSTD, and the main difficulties encountered are as follows:

- A large variation in size: Because of the large difference in the shooting distance, many targets usually vary in size from a few pixels to thousands. As a result, detection algorithms need to take into account the features of both smaller and larger targets.
- Irregular shape: Different targets at different distances can appear to have different shapes when looking at infrared images.
- Concealment: There are often noises such as clouds and buildings in the infrared image, which possess similar characteristics as small targets in specific cases. Their presence could affect the detection of targets.

The traditional IRSTD method is simple and has little randomness; different results have been collected using this method. For example, in the field of filtering, methods based on spatial domain filtering [1] and frequency domain filtering [2] can suppress and even out the background to some extent. Marvasti et al. [1] obtained an estimated image of the background of the infrared picture by filtering the background and then differentiating it from the original image to highlight small targets. This method is computationally efficient and responsive; however, it has limited accuracy. Wang et al. [2] demonstrated the frequency domain transformation of infrared images with non-negativity-constrained variational mode decomposition (NVMD) and enhanced small targets by performing frequency domain filtering. Compared with the null domain filtering method, this type of scheme increases the computational complexity, but significantly improves the accuracy. In the human vision

system (HVS) field, Hou et al. [3] published algorithms based on spectral residuals, which provide excellent results for small targets without a priori information and with simple texture and shape information, but have a lower ability to suppress complex backgrounds. Han et al. [4] proposed a multidirectional 2D least mean square (MDTDLMS) and a ratio difference joint local contrast measure (RDLCM) based on the local contrast method. They solved the problems of a small local contrast reference range and the insufficient effect of the differential ratio form algorithm based on a derivative entropy-based contrast measure (DECM) and obtained better detection results. However, robustness to dim targets was still lacking. In addition, Li et al. [5] proposed a detection method based on low-rank representation (LRR). This method uses LRR to decompose the infrared image into background and target components and completes the detection task by performing thresholding. This method has a better detection effect when the environment changes rapidly, it also has a high detection performance under strong noise background conditions. Although the above traditional methods can provide superior performance in some simple scenes, they still demonstrate shortcomings in terms of the segmentation ability of some complex cases, such as those with a complex background and strong noise. The reason behinds it is that the traditional approaches depend heavily on hyper-parametric adjustments and manually crafted features with limited expressiveness.

In recent years, in order to resolve the issues with traditional IRSTD methods,improve the accuracy of image detection, many of CNN-based methods have been implemented [6,7]. Dai et al. [8] proposed asymmetric contextual modulation (ACM) to enhance the target detection of classical semantic segmentation networks by combining contextual features in the model. This was the first time that semantic segmentation networks were applied in the field of infrared small target detection. ACMNet integrates a class-agnostic three-branch attention module to simultaneously measure the likelihood of each point in time containing action instances, action contexts, and non-action context frames. With the aim of collecting surface features to obtain detailed features, Zhang et al. [9] proposed a different structure calles feature compensation and cross-level correlation network (FC3-Net) to improve the encoder and decoder. Inspired by Taylor finite difference (TFD), Zhang et al. [10] incorporated the shape features of images in the network with success and developed an infrared shape network (ISNet) for IRSTD. Subsequently, Zhang et al. [11] presented a vision transformer called Runge-Kutta transformer (RKformer) for the first time in the IRSTD domain. It comes with a lightweight random-connection attention (RCA) module that learns sparse attention through random connections. Li et al. [12] proposed a dense nested attention network (DNANet), which combines the UNets of different depths by densely nesting interaction modules in three directions, which significantly improved the detection results of the model for small targets. A dense nested interaction module and a channel-space attention module are proposed to achieve level-by-level feature fusion and adaptive feature enhancement, and a new dataset is developed. However, as the above networks focus on discovering the deeper features of the image, they do not pay attention to details such as edges, which exist at the shallower levels of images. This results in the networks being able to roughly identify the location of the targets and determine the approximate range. These methods have proven to have good performance results in the detection of small infrared targets. however, the recognition of small targets such as the shape and other detailed features is lacking. If we use the edge information of the image well and apply it on infrared detection of small targets, we will get better performance.The utilization of edge features has been found to be effective in some fields. For example, Guo et al. [13] solved the problem of insufficient detail in synthetic aperture radar by using edge-preserving convolution, which decomposes the input into texture or content components. Correctly detecting the shape of the targets can provide the basis for further work, although this is still a challenging task in IRSTD.

To solve the problem of blurred contours in detection, we attempt to leverage the image curvature feature so that the model is able to account for the edges of the target, we also propose a novel curvature half-level fusion network (CHFNet) in this paper. First,

to ensure that the segmentation network interacts better with the image curvature feature, we designed a new bottom-up cross-layer feature fusion block, i.e., a half-level fusion (HLF) block. It resolves the issue of feature distortion at all levels of the network by extracting half-level features between two layers to increase the density of the feature level. We then propose the curvature attention mechanism. The mechanism first extracts the weighted mean curvature (WMC) [14] of the image as the curvature feature. Then, the feature matrix of the HLF blocks from different levels will interact with the curvature feature through the gated convolutional layer. The output of this stream is a matrix that contains the curvature feature and contextual information about the segmentation network. We refer to this as the curvature attention. Since the curvature algorithm mathematically combines first-order gradients and second-order gradients reliably, curvature attention leads to a finer extraction of the edges, thereby providing more detailed features in the model prediction. Finally, the output of the curvature attention branch is combined with the output of the segmentation network to achieve the mask prediction. Because of the curvature attention, the mask prediction of CHFNet is more complete and accurate, with a shape that is similar to the ground truth. Experimental results demonstrated that the proposed CHFNet outperformed the sate-of-the-art (SOTA) in the intersection over union (IoU), the normalized intersection over union (nIoU), and the probability of detection (P_d) features on the NUAA-SIRST dataset. Since the proposed model collects curvature information, it can provide a more complete mask prediction.

In summary, the contributions of this paper are as follows:

- We propose the novel CHFNet for IRSTD. The experiments on the public NUAA-SIRST database illustrated the effectiveness and robustness of the proposed CHFNet.
- We developed a curvature attention mechanism based on the curvature information of the image, which more reliably extracts the small target shape features, while suppressing complex background clutter to a certain extent.
- We designed a half-level fusion block, a new bottom-up cross-layer feature fusion method that minimizes the distortion at different levels of the feature.

The paper is structured as follows: First, we briefly review the related work and describe the architecture of each part of our CHFNet in detail in Section 2. Section 3 presents the experimental results, and Section 4 discusses the ablation study of the curvature. Conclusions are provided in Section 5.

2. Materials and Methods

2.1. Related Work

2.1.1. Infrared Small Target Detection

Traditional infrared small target detection methods are mainly based on various image processing algorithms that are computationally friendly, do not require learning, and have low randomness [15–17]. Among filter-based methods, there exist methods based on spatial domain filtering [1] and frequency domain filtering [2]. In terms of HVS-based methods, there exist algorithms based on spectral residuals [3] and local contrast [4]. In addition, some methods based on low-rank sparse matrices [5] have also been proposed. These traditional algorithms have a certain level of accuracy and are the basis for IRSTD. However, the robustness of traditional methods is still insufficient in terms of processing infrared images with complex background noise and faint targets.

In order to solve the problems of traditional methods, deep learning methods based on CNN networks [18,19] have been introduced for IRSTD. Zhao et al. [20] proposed TBC-net by building a target extraction module (TEM) and a semantic constraint module (SCM) to introduce deep learning methods to IRSTD for the first time and provided a method for synthesizing data. Zhao et al. [21] developed a prediction method based on a generative adversarial network (GAN), and used the IoU as a metric. Wang et al. [22] applied a cGAN network with a separate model building for missed detections (MDs) and false alarms (FAs) as two subtasks used as generators; then, they applied a discriminator for image classification to discriminate the output of the two generators and the three types of images

of the ground truth, and used model adversarial training to complete segmentation of images into objects or backgrounds. In [8], Dai et al. designed an asymmetric contextual modulation block. That is, two types of attention modules—bottom-up and top-down—were established. A complete network was constructed with a backbone based on FPNet and UNet. ACMNet was the first attempt to introduce semantic segmentation into infrared small targets, and has achieved good results in image recognition and image extraction compared to traditional methods. Subsequently, Dai et al. [23] presented the feature cyclic shift scheme mechanism to implement local contrast measurement and leveraged it to enhance ACMNet. The model solves the problem of insufficient detection of small targets in the direction of infrared detection targets better, with significant performance improvements compared to traditional methods. Zuo et al. [24] also proposed AFFPNet, which consists of unique feature extraction and feature fusion modules; they were able to achieve good results. Li et al. [12] proposed a dense nested attention network (DNANet) to improve the accuracy of the network for feature extraction by nesting UNets of different depths and facilitating three directions of interaction between the nodes. Lv et al. [25] proposed another attention nesting scheme using an asymmetric pyramid structure to achieve cross-layer feature fusion. To improve the signal-to-noise ratio of small targets while greatly reducing the false alarm rate, Zhang et al. [9] designed a cross-level feature correlation (CFC) module. Firstly, they utilized high-level features to build an energy filtering kernel. And then they leveraged it to suppress most of the background clutter in the mid-level features. Finally, pure target features are used to augment the targets in the output feature map. With mathematical interpretation, Zhang et al. [10] developed a Taylor finite difference (TFD)-inspired block to extract shape information and proposed a two-orientation attention aggregation (TOAA) block to calculate the low-level information in row and column directions. But the above CNN-based methods only stack convolutional layers to enlarge the receptive field, which are not efficient in extracting the global features. And these features are significant for the small target detection. Accordingly, Zhang et al. [11] presented for the first time the RKformer in the IRSTD domain. They designed a lightweight RCA module instead of the primitive self-attention in transformer for learning sparse attention through random connections.

Although the deep learning methods [26–28] have achieved excellent accuracy so far, there is still limited exploitation of the curvature feature in infrared images [29,30]. If we combine these methods with a more reliable curvature feature, the edge information of targets can be better detected. In this paper, we explored the role of curvature features in infrared images and designed a new cross-layer feature fusion mechanism. The network can be used to obtain more accurate edge features of the target by obtaining the curvature feature of the image, which provides a basis for extending the work such as by determining the type of target.

2.1.2. Cross-Layer Feature Fusion

As the most-effective semantic segmentation network, the CNN can process images at different levels with the convolutional layer and then combine them to obtain the final segmentation result [31,32]. Traditional cross-layer feature fusion networks include UNet [33], PANNet [34], FPNNet [35], etc. Redmon et al. [36] introduced cross-layer feature fusion in target detection and achieved superior results. In the field of IRSTD, Dai et al. proposed ACM [8] and ALC [23] as cross-layer feature fusion blocks based on contextual attention, and Li et al. [12] achieved the fusion of features multiple times for each level using a dense nested UNet. The fundamental purpose of cross-layer feature fusion is to reduce the distortion of features at each level while ensuring that the features of different levels are filtered with each other. To this end, we propose a novel half-level feature fusion method as a bottom-up cross-layer feature fusion mechanism to achieve minimal feature distortion at different levels.

2.1.3. Curvature-Based Image Processing

Curvature-related methods have been widely used in the image processing field [37]. The main forms of curvature that are widely used are Gaussian curvature, mean curvature, and differential curvature. Lee et al. [38] introduced Gaussian curvature in image filtering. Zhu et al. [39] employed the mean curvature to solve the image deblurring problem. Chen et al. [40] applied a differential curvature for image denoising. However, the traditional curvature requires complex nonlinear computation after multiple convolutions, which leads to its slow speed and poor real-time performance. Weighted mean curvature (WMC), proposed by Gong et al. [14], is a linear method with a faster computational speed. Specifically, WMC uses eight half-window Laplace operators to calculate the mean curvature in each of the eight directions and selects the minimum value among the eight channels to obtain the curvature information of the image. In this IRSTD method, we chose to apply WMC to process the image in order to extract the curvature feature in real-time and establish the attention mechanism.

2.2. Method

2.2.1. Overall Architecture

The overall architecture of the curvature half-level fusion network (CHFNet) is shown in Figure 1. In one branch, a single infrared image is fed into the stem block and encoded by the residual blocks to obtain the feature matrices x_0, x_1, x_2, and x_3 at different levels. After that, each feature matrix is fused across levels through the HLF blocks in a bottom-up direction. In the other branch, the curvature feature X_C of the image is extracted with WMC, and gated convolution is performed with the feature matrices to fuse X_C with the features of each level, resulting in the curvature attention X_A. We combined the output of the two branches in the form of attention and leveraged the convolutional segmentation head to perform fine segmentation of the blurred shapes to obtain the final mask prediction.

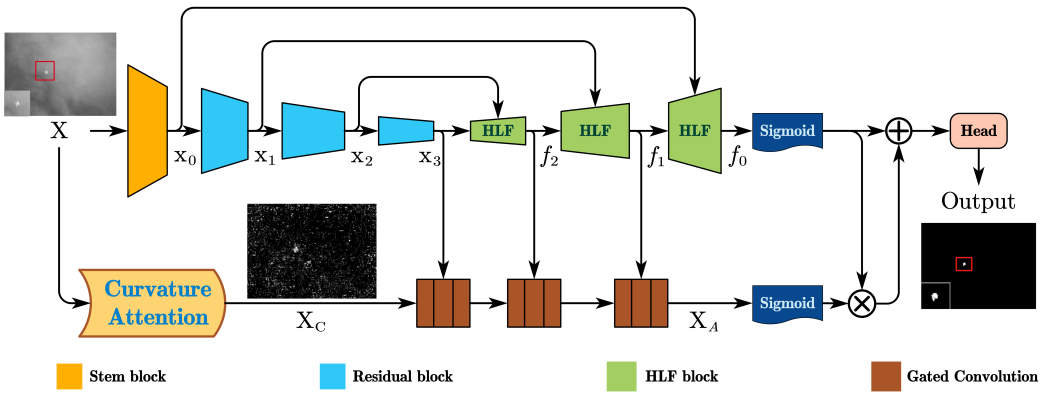


Figure 1. Overall architecture of CHFNet, with a backbone based on 4 depth UNets, HLF blocks, and curvature attention.

2.2.2. Half-Level Fusion Block

When performing cross-level feature fusion, we expected that the lower-level and higher-level feature will not be lost. With the aim to retain the effective features, we propose the half-level fusion (HLF) block. We assumed that the lower-level feature matrix is x with c_1 channels and the higher-level feature matrix is y with c_2 channels. It is clear that c_1 is greater than c_2. According to the structure of the overall network, the output matrix, F, of the HLF block has the same number of channels as the higher-level matrix, y. F is obtained for the HLF block by summing the four feature matrices with c_2 channels.

$$F = y + f_1 + f_2 + f_3, \tag{1}$$

where y is the higher level, i.e., representing the feature from a higher level. f_3 is obtained from the lower-level input x after upsampling, and the number of channels is transformed by the 1×1 convolutional layer, representing the feature from the lower level.

$$f_3 = V_1(U(x)), \tag{2}$$

where $V_i(\cdot)$ is the convolution with an $i \times i$ kernel and $U(\cdot)$ is upsampling. f_1, f_2 are obtained when the z matrix passes through different convolutional layers. Additionally, the z matrix is obtained by concatenating matrices x and y, which contain all features from the higher level and lower level, as shown in Equation (3).

$$z = C(U(x), y), \tag{3}$$

where $C(\cdot, \cdot)$ is the concat of two matrices in the channel dimension. Next, z passes through two convolutional layers of 1×1, respectively, and matrices f_1 with c_2 channels and z_1 with $c_2/2$ channels are obtained.

$$\begin{aligned} f_1 &= V_{1,1}(z) \\ z_1 &= V_{1,2}(z), \end{aligned} \tag{4}$$

where z is a matrix with $c_1 + c_2$ channels, f_1 and z_1 are attenuated in terms of information content compared to z, and $V_{1,1}$ and $V_{1,2}$ are two convolutional layers with 1×1 kernels. This means that $V_{1,1}$ and $V_{1,2}$ assume the function of selecting half of the features from the higher levels and lower levels, respectively. We refer to the middle-level formed by filtering half of the features from the higher and lower levels as the half-level. The feature matrices f_1 and f_2 that are obtained by the half-level are half-level features. f_1 is the z matrix obtained by performing only one convolution, which represents the more superficial part of the half-level feature. In order to avoid distortion during half-level filtering of the features, we used the z matrix for deeper features. The matrices z_{21} and z_{22} are the deeper features generated by the z matrix, which both have $c_2/2$ channels.

$$\begin{aligned} z_{21} &= V_3(V_1(z_1)) \\ z_{22} &= S(V_1(z_1)), \end{aligned} \tag{5}$$

where $S(\cdot)$ is a stream consisting of two sets of 3×3 convolutional layers, batch norm, and ReLU activation functions. z_{21} is obtained by z_1 passing the 3×3 convolutional layer, and z_{21} has $c_2/2$ channels. z_{22} is a matrix obtained by z_1 passing the stream S and also has $c_2/2$ channels. z_{21} and z_{22} are concatenated and integrated using convolution with a 1×1 kernel to derive the half-level feature, f_2, representing the deeper level.

$$f_2 = V_1(C(z_{21}, z_{22})) \tag{6}$$

At this point, we have obtained the four feature matrices that make up the output F, thereby completing the construction of the HLF block. The structure is shown in Figure 2.

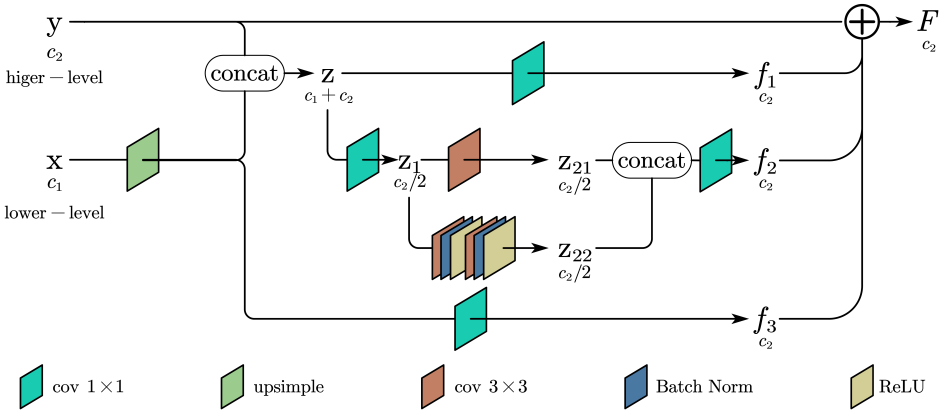


Figure 2. Structure of half-level fusion block.

2.2.3. Curvature Attention

Using grayscale images, we found that the targets retain certain edge features even in the blurred images. Traditional edge extraction methods generally apply first-order gradients of an image, while curvature consists of the first-order and second-order gradients combined in a mathematically provable way. Weighted mean curvature (WMC) methods [14] have the advantages of regularization, scalability, and a lower computational complexity; furthermore, we can construct branches for WMC in images, so as to extract the target feature and eliminate the effect of complex background clutter.

For any infrared image U, its WMC $H^w(U)$ is defined as shown in Equation (7).

$$H^w(U) = 2\|U\|_2 H(U) = \Delta U - \frac{U_x^2 U_{xx} + 2U_x U_y U_{xy} + U_y^2 U_{yy}}{U_x^2 + U_y^2}, \tag{7}$$

where Δ is the isotropic Laplace operator. U_x, U_y are the first-order gradients of the image in the horizontal and vertical directions, and U_{xx}, U_{yy}, U_{xy} are the second-order gradients of the image in different directions. $H(U)$ is the mean curvature of the image U, which is generally defined as Equation (8).

$$H(U) = \frac{(1 + U_x^2)U_{yy} - 2U_x U_y U_{xy} + (1 + U_y^2)U_{xx}}{2\sqrt{(1 + U_x^2 + U_y^2)^3}} \tag{8}$$

In the following section, we briefly describe the more efficient linear convolution method for the weighted mean curvature. First, we determine an isotropic Laplace operator k [41] as follows:

$$k = \begin{bmatrix} \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{6} & -1 & \frac{1}{6} \\ \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \end{bmatrix} \tag{9}$$

We derive eight half-window kernels to compute each directional normal curvature as follows:

$$\begin{aligned} k_1 &= \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{1}{3} & -1 & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 \end{bmatrix} \quad k_2 = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & -1 & \frac{1}{6} \\ 0 & 0 & 0 \end{bmatrix} \quad k_3 = \begin{bmatrix} 0 & \frac{1}{6} & \frac{1}{6} \\ 0 & -1 & \frac{1}{3} \\ 0 & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \quad k_4 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{6} & -1 & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{bmatrix} \\ k_5 &= \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{12} \\ \frac{1}{3} & -1 & 0 \\ \frac{1}{12} & 0 & 0 \end{bmatrix} \quad k_6 = \begin{bmatrix} \frac{1}{12} & \frac{1}{3} & \frac{1}{6} \\ 0 & -1 & \frac{1}{3} \\ 0 & 0 & \frac{1}{12} \end{bmatrix} \quad k_7 = \begin{bmatrix} 0 & 0 & \frac{1}{12} \\ 0 & -1 & \frac{1}{3} \\ \frac{1}{12} & \frac{1}{3} & \frac{1}{6} \end{bmatrix} \quad k_8 = \begin{bmatrix} \frac{1}{12} & 0 & 0 \\ \frac{1}{3} & -1 & 0 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{12} \end{bmatrix} \end{aligned} \tag{10}$$

We convolved each channel of the original image U with each of the eight kernels and stitched the resulting eight matrices in the dimension of the channels into a new matrix as a union of them. Since the curvature of the targets is generally prominent in

all directions, we selected the smallest absolute value for each pixel in the eight channels as the final output pixel. From this, we combined the eight channels into one channel and then combined the results of each channel in U, keeping the number of input and output channels the same. At this point, we obtained the curvature feature X_C of the image, with some attenuation of the noisy pixels.

$$X_C = H^w(X), \tag{11}$$

where X is the input infrared image. In the following, we fused the curvature features of the image with the features of each level by performing a gated convolution to form the final curvature attention. Assuming that the gated convolution is $g(x, y)$, we specify the gated convolution stream $G(c, x, y, z)$ as

$$G(c, x, y, z) = g(g(g(c, x), y), z), \tag{12}$$

where c is the curvature feature of the image and x, y, z is the feature matrix of different levels in the semantic segmentation network. By inserting Equation (12) into CHFNet, we obtain the relationship between the curvature attention X_A and the curvature feature X_C, as well as the feature matrices x_3, f_2, f_1 of each level.

$$X_A = G(X_C, x_3, f_2, f_1) \tag{13}$$

The image in the final curvature attention is shown in the right of Figure 3. Since the original image only acts in the model, its grayscale value is too low to be observed by humans. To better observe the effect of the curvature attention mechanism on the initial curvature image, we zoomed in on the elements of the original image 1000 times. In this case, except for the higher grayscale values around the two target locations, all areas remained at lower grayscale values even after the number of runs increased by 1000. It can be seen that, by performing gated convolution, the model can restrict the attention to the vicinity of the targets. After obtaining the curvature attention X_A, it will be combined with the output of the semantic segmentation network f_0. First, we activated the two matrices with the sigmoid function to make the prediction more obvious.

$$\begin{aligned} F &= sigmoid(f_0) \\ A &= sigmoid(X_A) \end{aligned} \tag{14}$$

Subsequently, we combined F and A in the form of attention, and then, the final output of the model is obtained by processing the segmentation head.

$$Output = head(F \otimes A + F), \tag{15}$$

where $head(\cdot)$ is the segmentation head and $\otimes$ is elementwise multiplication.

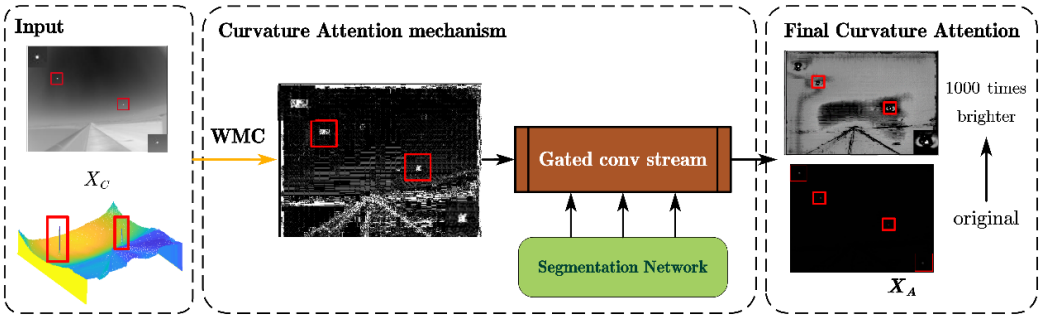


Figure 3. Structure of curvature attention branch.

2.2.4. Loss Function

The IoU is an important measure of model accuracy in semantic segmentation, and the IoU loss is a common loss function defined according to the IoU to measure the difference between the mask prediction and the ground truth. It has stronger convergence compared to the L_2 loss function. The IoU loss is defined as follows:

$$L_{IoU} = -ln\frac{area_{inter}}{area_{all}}, \tag{16}$$

where $area_{inter}$ is the number of pixels in the intersection part of the mask prediction and the ground truth and $area_{all}$ is the number of pixels in the union part of the mask prediction and the ground truth.

3. Results

3.1. Evaluation Metrics

(1) Intersection over union: We leveraged the common intersection over union (IoU) as one of the training metrics. The IoU uses the fraction of correct predictions $area_{inter}$ in the union of predictions and the ground truth $area_{all}$ to measure the prediction accuracy of the model.

$$IoU = \frac{area_{inter}}{area_{all}} \tag{17}$$

(2) Normalized intersection over union: According to [8], we used normalized intersection over union for the NUAA-SIRST dataset. The nIoU is a metric obtained by normalizing the IoU score of each item in the dataset, which is better for measuring the performance of the model in each item of the data.

$$nIoU = \frac{1}{n}\frac{area_{inter}}{area_{all}} \tag{18}$$

In addition, Li et al. [12] proposed new metrics for evaluating the quality of infrared small target detection, the probability of detection (P_d) and the false alarm rate (F_a). These two metrics have also become the mainstream metrics for evaluating the quality of detection.

(3) Probability of detection: This is the ratio of the number of pixels of the correct prediction part $area_{inter}$ and the number of pixels of the ground truth $area_{ground}$.

$$P_d = \frac{area_{inter}}{area_{ground}} \tag{19}$$

(4) False alarm rate: This is the ratio of the number of pixels in the wrong prediction part $area_{false}$ to the number of pixels in all the images $area_{image}$.

$$F_a = \frac{area_{false}}{area_{image}} \tag{20}$$

We applied these four metrics to measure the performance of the proposed CHFNet on the public NUAA-SIRST benchmark.

3.2. Experiment Settings and Dataset

Our training optimizer was Adam; the learning rate was set to 0.001; the epochs were 800; the batch size was 8. All models were implemented in PyTorch on a computer with a 24 vCPU Intel Xeon Platinum 8255C CPU @ 2.50 GHz CPU and an RTX 3090 GPU. We conducted the experiment on the NUAA-SIRST dataset, which is currently widely used. The images of the dataset contain both single and multiple targets, and the small target occupies a very small portion of the image space. The CNN-based IRSTD methods we compared were MDvsFA [22], ACMNet [8], ALCNet [23], and DNANet [12]; the traditional algorithms we selected were TopHat [42], max-median [43], WSLCM [44], TLLCM [45], IPI [46], NRAM [47], RIPT [48], PSTNN [49], and MSLSTIPT [50].

3.3. Experimental Results and Comparison

The value of studying some of the traditional methods is more limited because the metrics of the traditional methods are quite different from the deep learning methods. Because of this, we chose the TopHat and IPI methods among the traditional methods. TopHat is the classical method in target detection, which determines the target position according to the difference between the open operation and the original image. This approach works better for complex backgrounds in IR images; therefore, it was widely used in early IR small target problems. Additionally, IPI is one of the more promising metrics among all traditional methods.

Among the currently common algorithmic models, we selected MDvsFA, ACMNet, ALCNet, and DNANet. MDvsFA is optimized using adversarial networks due to its special problem-solving approach, and the structure becomes simpler. It optimized FA and MD separately and lets them balance by adversarial learning, which reduces the complexity of previous networks and increases the flexibility of sub-task selection networks to achieve excellent performance. ACMNet is a more advanced approach that combines attention modules with contextual branches and does not focus only on the current action module, which outperforms traditional state-of-the-art methods. ALCNet is proposed for situations where small target feature information is easily lost. The modified algorithm is based on the method of cyclic displacement of feature maps, and design a set of acceleration strategies. The DNANet method designs densely nested modules to interact with the upper and lower layers of information to make the obtained target features clearer. It uses U-Net as the basic network structure and keeps increasing the level of its network to obtain deeper semantic information and a larger perceptual field.

However, due to the simple structure of MDvsFA, it cannot obtain as high of an accuracy as the other three methods, and has the worst performance among all the exemplified advanced algorithms, but still has some reference significance. The ACMNet and ALCNet methods are more advanced. They have similar performance, and since ALCNet is an enhanced algorithm of ACMNet. However, they do not perform as well as DNANet in terms of obtaining a clear target due to the use of dense nesting. It is more robust to changes in clutter background, target size and target shape, but the ability to extract edge information is insufficient.

We designed CHFNet to improve the design method so as to retain internal information with excellent results and to also propose an innovative method that incorporates curvature attention to enhance the detection of small target edge information. Our method performed excellently for all metrics.

As shown in Figure 4, we plotted the ROC curves of different methods on NUAA-SIRST. It can be seen that our model performed better than the other methods. Additionally, the area under the ROC curve (AUC) of the proposed CHFNet was also significantly better than that of the other methods. For example, the AUC of CHFNet reached 0.9544, while the AUC of DNANet was 0.9461.

As shown in Table 1, We calculated the IoU, nIoU, P_d, fa of all the enumerated methods for the NUAA-SIRST dataset, and in these data we can clearly see that the traditional methods are less effective and the enumerated advanced algorithms have a great improvement in performance. The IoU of IPI has improved and the F_a has also decreased. However, these traditional algorithms perform relatively poorly in terms of performance. Among the advanced algorithms we cite, MDvsFA has an IoU of 60.30, which is poor in comparison. acMNet is 74.81, ALCNet is 74.31, both with similar performance, and DNANet is 70.04. The performance of these advanced algorithms has a great improvement compared to the traditional algorithms, and the values of IoU, nIoU, and P_d are highly improved, and the value of F_a has significant reduction. These are all high performance performances. Our CHFNet has the best results among all the algorithms, where the IoU of 78.76 is greater than all the listed methods, the P_d of 98.91 is the best performance among all the methods, and the value of F_a of 1.814 is better than expected.

Table 1. Comparison of the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}) for different detection methods on NUAA-SIRST.

Model	IoU	nIoU	P_d	F_a	Model	IoU	nIoU	P_d	F_a
TopHat	7.143	5.201	79.84	101.2	PSTNN	22.40	22.35	77.95	2.911
Max-Median	4.172	2.150	69.20	5.533	MSLSTIPT	10.30	9.58	82.13	113.1
WSLCM	1.158	0.849	77.95	544.6	MDvsFA	60.30	58.26	89.35	5.635
TLLCM	1.029	0.905	79.09	589.9	ACMNet	74.81	75.09	97.95	1.024
IPI	25.67	24.57	85.55	1.147	ALCNet	74.31	73.12	97.34	2.021
NRAM	12.16	10.22	74.52	1.385	DNANet	70.04	69.45	95.44	3.073
RIPT	11.05	10.15	79.08	2.261	**CHFNet**	**78.76**	**77.65**	**98.91**	**1.814**

By comparing different algorithms in ROC curves, and comparing different algorithms in tabular data, we can visualize the advantages of advanced algorithms over traditional algorithms in terms of accuracy and precision of target detection. We can also see that, compared with other advanced algorithms, we use CHFNet, introduce curvature attention, enhance the calculation of target edge information, and detect the target edge information more accurately, so that our method has a superior performance among all methods, and the detection of targets is more accurate

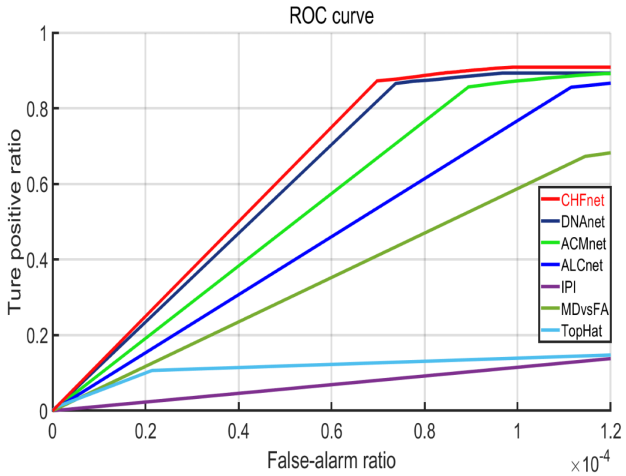


Figure 4. ROC curve of different methods.

3.4. Visual Results

As shown in Figure 5, we visualized the experimental results of the method using the ROC curves, as shown in Figure 4. Since we leveraged the HLF block and curvature attention to help reconstruct the shape of the target, CHFNet had the edge feature of the mask prediction that was closer to the ground truth and had better overall predictions than the other methods.

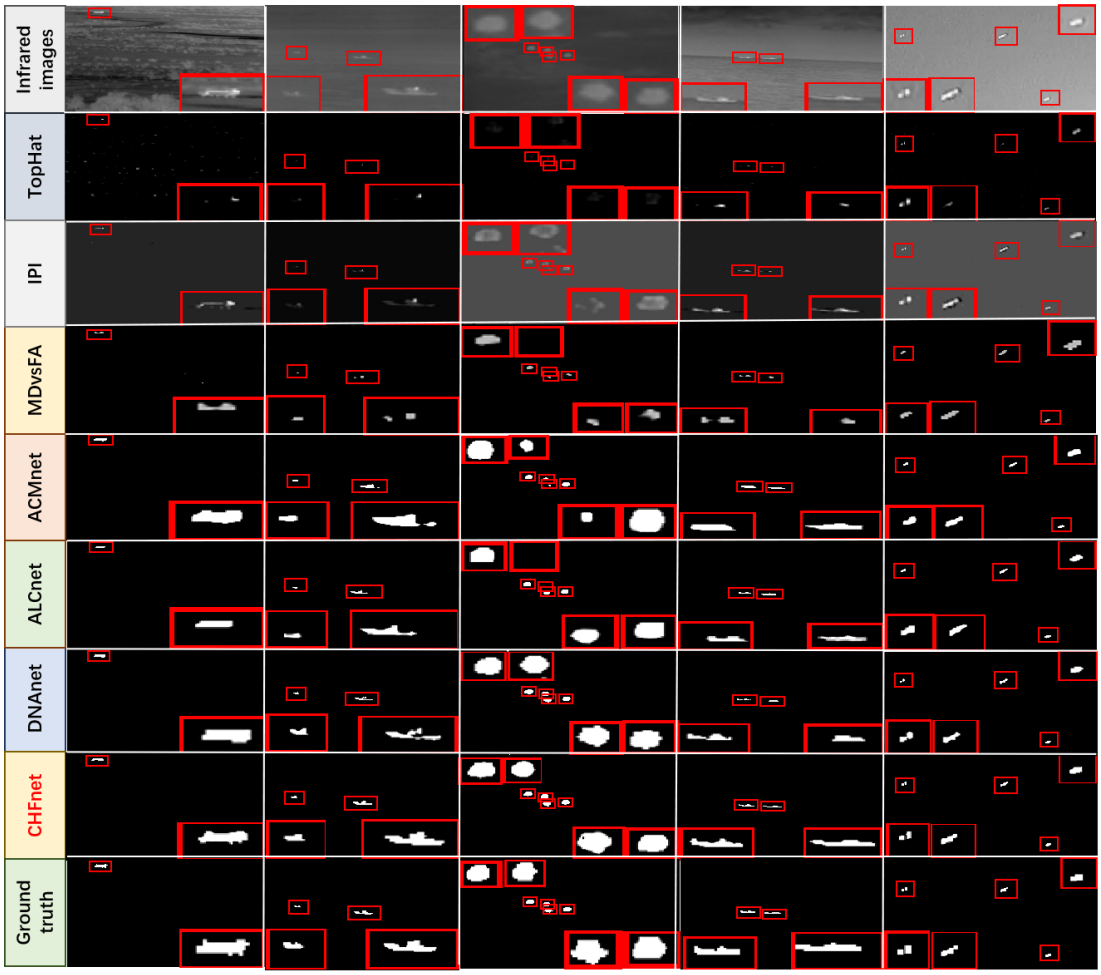


Figure 5. Visual results of different methods for IRSTD. Red boxes are the location of the targets. Closed-up views are shown.

3.5. Ablation Study

Ablation study of different innovation components: As shown in Table 2, we ablated the curvature attention and HLF block. In the results, we found that only using UNet led to poor segmenting, an inability to extract the complete feature of the target, and sensitivity to clutter, meaning that both P_d and F_a performed poorly. Additionally, for UNet and curvature attention fusion, we found that the curvature attention can improve the nIoU metric well; however, F_a was high due to the insufficient constraint of the segmentation feature in UNet on the curvature feature. In the fusion experiment of UNet and the HLF block, we found that the IoU of the HLF block for mask prediction improvement was sufficient. The fusion structure of the two components achieved complementary performance and improved the explored metrics.

Table 2. Ablation study of different innovation components for the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}).

Method	IoU	nIoU	P_d	F_a
UNet	69.28	71.77	94.50	2.653
UNet + curvature	75.80	76.35	97.91	3.402
UNet + HLF	77.55	75.38	98.80	0.138
HLF + curvature	78.76	77.65	98.91	1.814

Impact of HLF block: We studied the impact of the HLF blocks on the network, as shown in Table 3, where n denotes the number of HLF blocks in the network. Levels

without HLF blocks were replaced with the classic UNet decoder. When there was no HLF block in the network, the model was susceptible to disturbance by clutter, meaning that the predictions were poor. With the increasing number of HLF blocks, the prediction results of the network were gradually optimized. This indicated that the HLF block was important to the prediction quality.

Table 3. Ablation study of number of HLF blocks (n) for the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}).

n	IoU	nIoU	P_d	F_a
0	75.80	76.35	97.90	3.402
1	76.11	76.47	97.96	4.280
2	76.89	77.52	98.17	2.129
3	78.76	77.65	98.91	1.814

Impact of gated convolutional layer in curvature attention: We studied the impact of the gated convolutional layer (GCL) in the curvature attention branch, as shown in Table 4, where n denotes the number of GCLs. When the curvature feature was directly used as the attention mechanism during the interactions with the segmentation network, P_d and F_a were low. As the GCL was gradually added to the curvature attention branch, the F_a of the model tended to decrease, and the metrics such as the IoU, nIoU, and P_d, which measure the prediction accuracy of the model, gradually improved. This indicated that curvature attention through the GCL was important when interacting with the segmentation network.

Table 4. Ablation study of the number of GCLs (n) for the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}).

n	IoU	nIoU	P_d	F_a
0	76.37	74.95	97.77	0.528
1	76.42	74.70	98.52	3.087
2	77.35	75.63	98.65	2.852
3	78.76	77.65	98.91	1.814

4. Discussion

In this section, we discuss different curvature extraction methods for IRSTD. During image processing, a curvature is formed with different mathematical definitions. In order to compare the performance with CHFNet, we chose to compare WMC with the mean curvature (MC) and the difference curvature (DC) for experimental purposes.

In Table 5, we conducted different curvature experiments in relation to the following four metrics: IoU, nIoU, P_d, and F_a. We found that all metrics of the proposed CHFNet with WMC to extract the image curvature feature were significantly better than when other forms of curvature were used.

Table 5. Different curvature experiment for the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}).

Curvature	IoU	nIoU	P_d	F_a
WMC	78.76	77.65	98.91	1.814
MC	74.40	73.50	98.80	2.391
DC	76.97	74.58	98.01	0.408

As shown in Figure 6, we compared the ROC curves of different curvatures including WMC, MC, and DC on the NUAA-SIRST database. It can be seen that the AUC of WMC was significantly better than the other forms of curvature. Thus, we can prove that there was performance superiority in using WMC in the proposed CHFNet.

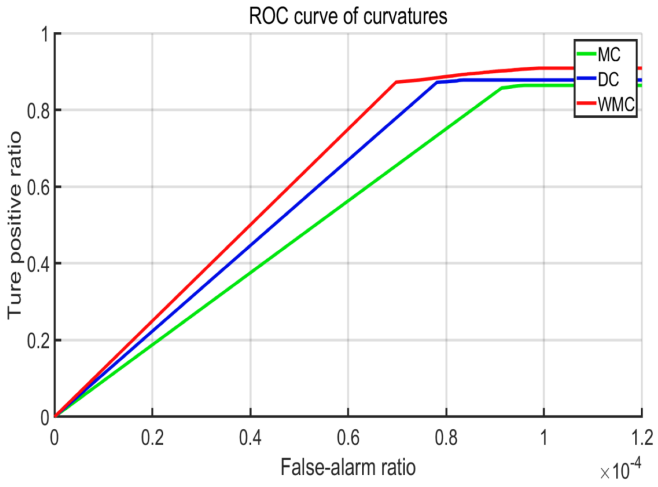


Figure 6. ROC curve of different curvatures.

In Figure 7, we demonstrate the curvature feature image extracted by different curvature forms. In order to extract the edge features of small targets, we should require the curvature to extract the shape of small targets as much as possible while suppressing the effect of background clutter. Among these methods, although DC extracted the edge features of small targets well, it failed to remove the background clutter. MC suppressed the background clutter, but removed the targets as a type of noise in some cases (e.g., the second to fourth columns). Only WMC took into account both requirements and led to better performance.

Furthermore, the image in the last column was obtained by increasing the mean value of the infrared image in the first column to Gaussian white noise, and the comparison between the noise image and the original image is shown in Figure 8. The mean value of the noise was 10^{-4}, and the variance was 8×10^{-6}. The curvature minimum value was 21.2500. It can be seen that all the curvatures were more sensitive to Gaussian noise, but the weighted mean curvature still maintained a relatively stable extraction effect.

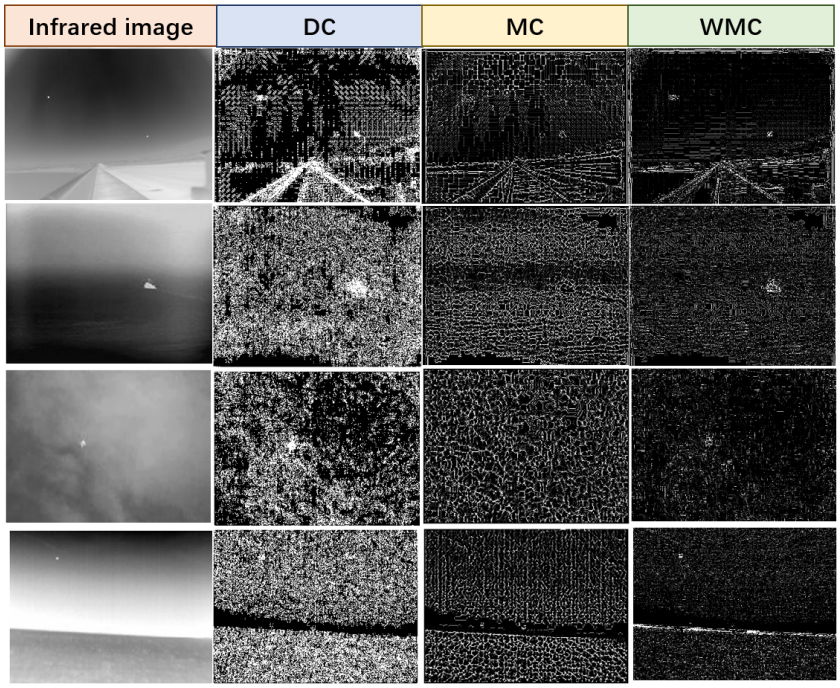


Figure 7. Visual results of different curvatures.

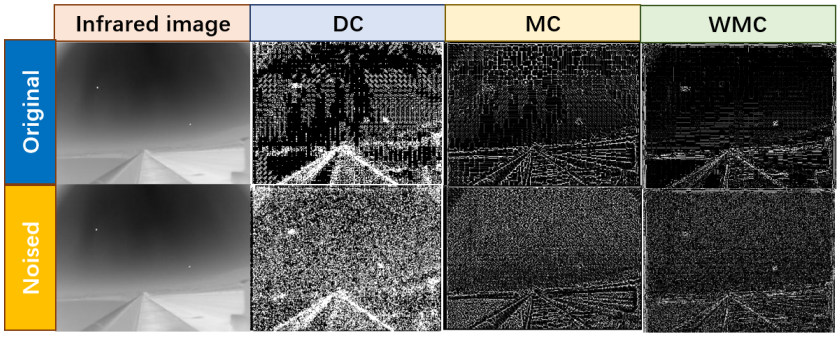


Figure 8. Comparison of noise image and original image.

In practical applications, the image will be disturbed by some noise. From Figure 8, we can see that curvature extraction was more sensitive to high-frequency noise, and filtering in the modeling pre-stage is a common measure to manage high-frequency noise. Additionally, different filtering methods can produce different levels of loss of detail in the image. In this experiment, we selected the mean filter and Gaussian filter commonly used in spatial filtering and the DCT filter commonly used in frequency domain filtering, and their filtering effects and details were compared, as shown in Figure 9; the experimental results are shown in Table 6.

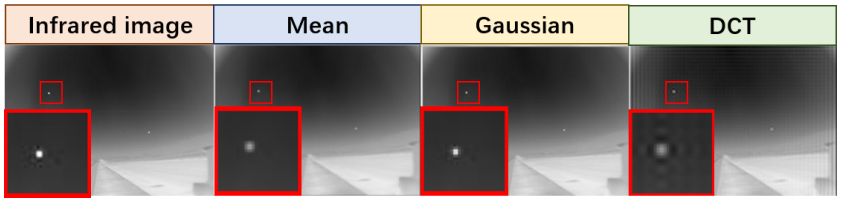


Figure 9. Comparison of filtering effects.

Table 6. Effect of filtering on model accuracy for the IoU (%), nIoU (%), P_d (%), and F_a (10^{-5}).

Indicators	Mean	Gaussian	DCT	No Filter
IoU	72.72	77.90	74.95	78.76
nIoU	73.88	75.54	73.75	77.65
P_d	98.08	98.08	98.08	98.91
F_a	5.052	0.634	2.315	1.814

As can be seen from the Table 6, these filtering methods had side effects on the curvature features of the model because the filtering denoising blurred some of the detailed features of the image and affected the extraction of these details by the curvature attention mechanism. However, among them, Gaussian filtering performed relatively well in terms of the IoU, nIoU, and F_a. This was probably due to the fact that a Gaussian distribution was used to determine the weights of the convolution kernel, thus allowing a higher level of information to be retained in the filtering results.

To further explore the detection capability of the curvature for small targets, we collected a batch of items with the smallest target size in NUAA-SIRST. The size of these small targets varied from 4 to 10 pixels, and we plotted the mean IoU of CHFNet's prediction results in Figure 10. As shown in the figure, the prediction results were more stable when the target size reached six pixels and above. However, the number of items with as small a target in the NUAA-SIRST dataset was limited, and the interference of the clouds, the ocean, and other factors in the scene of the model could not be avoided; therefore, this study was for reference only.

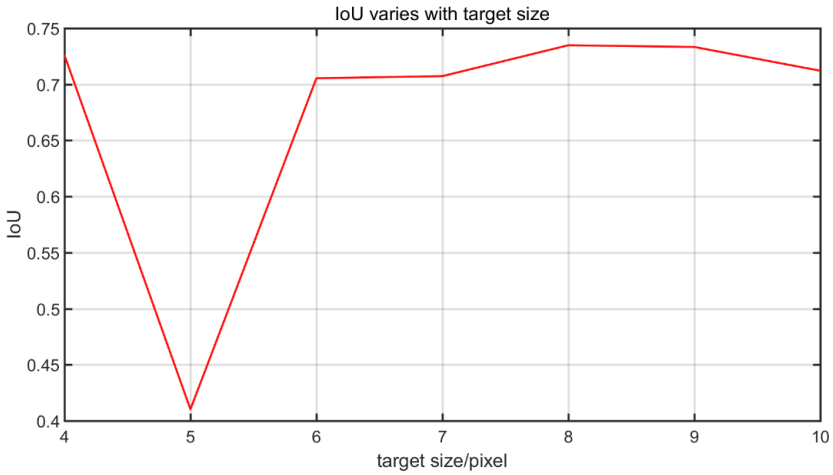


Figure 10. Curve of target pixel size versus the IoU.

5. Conclusions

In this paper, we proposed CHFNet for IRSTD. First, we designed an HLF block, which can be used to implement cross-layer feature fusion better, thus allowing the segmentation network to extract target features more precisely. Subsequently, we introduced the mean curvature feature to extract the edge information of the image. After obtaining the curvature feature of the image, we obtained the final curvature attention matrix by fusing them with the features of each layer of the segmentation network and filtering them together. The curvature attention was fused with the results of the segmentation network and then processed by the segmentation head to obtain the final mask prediction. The proposed CHFNet achieved considerable improvement in the performance on the public NUAA-SIRST dataset. Since the edge features of the target were extracted using the curvature attention feature, compared with other deep learning methods, the mask prediction of CHFNet can better retain the edge information of the target, which provides a basis for further work, such as determining the kind of target.

Author Contributions: Conceptualization, H.B.; Methodology, M.Z.; Software, B.L.; Resources, K.Y.; Writing—original draft, T.W.; Writing—review & editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grants 62272363, 62036007, 62061047, 62176195, and U21A20514, the Equipment Advance Research Field Fund Project under Grant 80913010601, the Young Elite Scientists Sponsorship Program by CAST under Grant 2021QNRC001, the Shaanxi Province Key Research and Development Program Project under Grant 2021GY-034, the Youth Talent Promotion Project of Shaanxi University Science and Technology Association under Grant 20200103.

Data Availability Statement: The NUAA-SIRST dataset is downloaded free of charge from the article according to the link https://github.com/YeRen123455/Infrared-Small-Target-Detection.

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

References

1. Marvasti, F.S.; Mosavi, M.R.; Nasiri, M. Flying small target detection in IR images based on adaptive toggle operator. *IET Comput. Vis.* **2018**, *12*, 527–534. [CrossRef]
2. Anju, T.S.; Raj, N.R.N. Shearlet transform based image denoising using histogram thresholding. In Proceedings of the 2016 International Conference on Communication Systems and Networks (ComNet), Thiruvananthapuram, India, 21–23 July 2016; pp. 162–166. [CrossRef]
3. Hou, X.; Zhang, L. Saliency Detection: A Spectral Residual Approach. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8. [CrossRef]
4. Wei, Y.; You, X.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **2016**, *58*, 216–226. [CrossRef]

5. Li, M.; He, Y.; Zhang, J. Small Infrared Target Detection Based on Low-Rank Representation. In Proceedings of the 8th International Conference, ICIG 2015, Tianjin, China, 13–16 August 2015. [CrossRef]
6. Zhang, M.; Wang, N.; Li, Y.; Gao, X. Deep latent low-rank representation for face sketch synthesis. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3109–3123. [CrossRef]
7. Zhang, M.; Li, Y.; Wang, N.; Chi, Y.; Gao, X. Cascaded face sketch synthesis under various illuminations. *IEEE Trans. Image Process.* **2020**, *29*, 1507–1521. [CrossRef]
8. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Asymmetric Contextual Modulation for Infrared Small Target Detection. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2021; pp. 949–958. [CrossRef]
9. Zhang, M.; Yue, K.; Zhang, J.; Li, Y.; Gao, X. Exploring Feature Compensation and Cross-Level Correlation for Infrared Small Target Detection. In Proceedings of the 30th ACM International Conference on Multimedia, MM'22, Lisboa, Portugal, 10–14 October 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1857–1865. [CrossRef]
10. Zhang, M.; Zhang, R.; Yang, Y.; Bai, H.; Zhang, J.; Guo, J. ISNet: Shape Matters for Infrared Small Target Detection. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 867–876. [CrossRef]
11. Zhang, M.; Bai, H.; Zhang, J.; Zhang, R.; Wang, C.; Guo, J.; Gao, X. RKformer: Runge-Kutta Transformer with Random-Connection Attention for Infrared Small Target Detection. In Proceedings of the 30th ACM International Conference on Multimedia, MM'22, Lisboa, Portugal, 10–14 October 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1730–1738. [CrossRef]
12. Ren, D.; Li, J.; Han, M.; Shu, M. DNANet: Dense Nested Attention Network for Single Image Dehazing. In Proceedings of the ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 2035–2039. [CrossRef]
13. Guo, J.; He, C.; Zhang, M.; Li, Y.; Gao, X.; Song, B. Edge-Preserving Convolutional Generative Adversarial Networks for SAR-to-Optical Image Translation. *Remote Sens.* **2021**, *13*, 3575. [CrossRef]
14. Gong, Y.; Goksel, O. Weighted mean curvature. *Signal Process.* **2019**, *164*, 329–339. [CrossRef]
15. Zhang, M.; Wang, N.; Gao, X.; Li, Y. Markov random neural fields for face sketch synthesis. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, Stockholm, Sweden, 13–19 July 2018; pp. 1142–1148. [CrossRef]
16. Zhang, M.; Wang, N.; Li, Y.; Wang, R.; Gao, X. Face sketch synthesis from coarse to fine.In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32. [CrossRef]
17. Zhang, M.; Li, J.; Wang, N.; Gao, X. Recognition of facial sketch styles. *Neurocomputing* **2015**, *149*, 1188–1197. [CrossRef]
18. Zhang, M.; Xin, J.; Zhang, J.; Tao, D.; Gao, X. Microscope Chip Image Super-Resolution Reconstruction via Curvature Consistent Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**.
19. Zhang, M.; Wu, Q.; Zhang, J.; Gao, X.; Guo, J.; Tao, D. Fluid micelle network for image super-resolution reconstruction. *IEEE Trans. Cybern.* **2022**, *53*, 578–591. [CrossRef]
20. Zhao, M.; Cheng, L.; Yang, X.; Feng, P.; Liu, L.; Wu, N. TBC-Net: A real-time detector for infrared small target detection using semantic constraint. *arXiv* **2019**, arXiv:2001.05852.
21. 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.* **2021**, *59*, 4481–4492. [CrossRef]
22. Wang, H.; Zhou, L.; Wang, L. Miss Detection vs. False Alarm: Adversarial Learning for Small Object Segmentation in Infrared Images. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8508–8517. [CrossRef]
23. 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]
24. Zuo, Z.; Tong, X.; Wei, J.; Su, S.; Wu, P.; Guo, R.; Sun, B. AFFPN: Attention Fusion Feature Pyramid Network for Small Infrared Target Detection. *Remote Sens.* **2022**, *14*, 3412. [CrossRef]
25. Lv, G.; Dong, L.; Liang, J.; Xu, W. Novel Asymmetric Pyramid Aggregation Network for Infrared Dim and Small Target Detection. *Remote Sens.* **2022**, *14*, 5643. [CrossRef]
26. Zhang, M.; Wang, N.; Li, Y.; Gao, X. Bionic face sketch generator. *IEEE Trans. Cybern.* **2019**, *50*, 2701–2714. [CrossRef]
27. Zhang, M.; Zhang, J.; Chi, Y.; Li, Y.; Wang, N.; Gao, X. Cross-domain face sketch synthesis. *IEEE Access* **2019**, *7*, 98866–98874. [CrossRef]
28. Zhang, M.; Wu, Q.; Guo, J.; Li, Y.; Gao, X. Heat transfer-inspired network for image super-resolution reconstruction. *IEEE Trans. Onneural Netw. Learn. Syst.* **2022**, 1–11. [CrossRef]
29. Zhang, M.; Wang, N.; Li, Y.; Gao, X. Neural probabilistic graphical model for face sketch synthesis. *IEEE Trans. Onneural Netw. Learn. Syst.* **2020**, *31*, 2623–2637. [CrossRef]
30. Zhang, M.; Li, J.; Wang, N.; Gao, X. Compositional model-based sketch generator in facial entertainment. *IEEE Trans. Cybern.* **2017**, *48*, 904–915. [CrossRef]
31. Zhang, S.; Gao, X.; Wang, N.; Li, J.; Zhang, M. Face sketch synthesis via sparse representation-based greedy search. *IEEE Trans. Image Process.* **2015**, *24*, 2466–2477. [CrossRef]
32. Zhang, M.; Wang, R.; Gao, X.; Li, J.; Tao, D. Dual-transfer face sketch–photo synthesis. *IEEE Trans. Image Process.* **2019**, *28*, 642–657. [CrossRef] [PubMed]

33. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505.04597.
34. Wang, W.; Xie, E.; Song, X.; Zang, Y.; Wang, W.; Lu, T.; Yu, G.; Shen, C. Efficient and Accurate Arbitrary-Shaped Text Detection with Pixel Aggregation Network. *arXiv* **2019**, arXiv:1908.05900.
35. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. *arXiv* **2016**, arXiv:1612.03144.
36. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv* **2015**, arXiv:1506.02640.
37. Zhang, M.; Xin, J.; Zhang, J.; Tao, D.; Gao, X. Curvature consistent network for microscope chip image super-resolution. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [CrossRef]
38. Lee, S.H.; Seo, J.K. Noise removal with Gauss curvature-driven diffusion. *IEEE Trans. Image Process.* **2005**, *14*, 904–909. [CrossRef]
39. Fairag, F.; Chen, K.; Ahmad, S. An effective algorithm for mean curvature-based image deblurring problem. *Comput. Appl. Math.* **2022**, *41*, 176. [CrossRef]
40. Chen, Q.; Montesinos, P.; Sun, Q.S.; Heng, P.A. Adaptive total variation denoising based on difference curvature. *Image Vis. Comput.* **2010**, *28*, 298–306. [CrossRef]
41. Kamgar-Parsi, B.; Kamgar-Parsi, B.; Rosenfeld, A. Optimally isotropic Laplacian operator. *IEEE Trans. Image Process.* **1999**, *8*, 1467–1472. [CrossRef]
42. Bai, X.; Zhou, F. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognit.* **2010**, *43*, 2145–2156. [CrossRef]
43. Deshpande, S.D.; Er, M.H.; Venkateswarlu, R.; Chan, P. Max-mean and max-median filters for detection of small targets. In *Signal Data Processing of Small Targets*; SPIE: Bellingham, WA, USA, 1999; Volume 3809, pp. 74–83. [CrossRef]
44. 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.* **2021**, *18*, 1670–1674. [CrossRef]
45. 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]
46. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared Patch-Image Model for Small Target Detection in a Single Image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef]
47. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared Small Target Detection via Non-Convex Rank Approximation Minimization Joint l2,1 Norm. *Remote Sens.* **2018**, *10*, 1821. [CrossRef]
48. 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]
49. Zhang, L.; Peng, Z. Infrared Small Target Detection Based on Partial Sum of the Tensor Nuclear Norm. *Remote Sens.* **2019**, *11*, 382. [CrossRef]
50. 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.* **2021**, *59*, 3737–3752. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.