IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 21, 2022, accepted May 5, 2022, date of publication May 10, 2022, date of current version May 19, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3174054

Dim Space Target Detection via Convolutional Neural Network in Single Optical Image

XIANGJI GUO[1,2], TAO CHEN[1], JUNCHI LIU[1], YUAN LIU[1,2], AND QICHANG AN[1]
[1]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
[2]College of Optoelectronics, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China

Corresponding author: Junchi Liu (liujunchi703@163.com)

This work was supported in part by the Natural Science Foundation of China under Grant 62005279, and in part by the Youth Innovation Promotion Association of the Chinese Academy of Sciences under Grant 2020221.

ABSTRACT Real-time dim space target detection is a significant challenge in space situation awareness. This paper proposes a single-frame space object segmentation and detection method based on deep learning. Firstly, Channel and Space Attention U-net (CSAU-Net) is presented based on space image features. We remove unnecessary feature layers and add attention modules in the traditional encoder and decoder structure to enhance feature fusion and better use original feature layers. The proposed network structure can achieve accurate segmentation of space objects with fewer data training. At the same time, we construct a space target dataset for training, which contains targets with different signal-to-noise ratios to enhance the generalization of convolutional neural networks. After obtaining the segmentation masks, a simple connected component labeling method is applied to extract the centroid of the space target. Experiments show that our approach can achieve an ideal segmentation effect when the signal-to-noise ratio of the space target in the simulated dataset is 0.3. In addition, the proposed algorithm can realize fast segmentation and achieve an accuracy of 98.5%, which is similar to the traditional multi-frame space target detection method in real space image detection.

INDEX TERMS Convolutional neural networks, semantic segmentation, dim space target detection, connected component labeling.

I. INTRODUCTION

With the rapid development of space technology, space objects in earth orbit increased dramatically during the past few decades [1]. Break-ups, explosions, and collisions produced a mass of space debris [2]. This debris will seriously affect the routine flight of space aircraft and satellites. It will cause damage to satellites mission failure and bring significant Loss to space engineering when space debris collides with a spacecraft. The trash from the collision can also threaten spacecraft. Therefore, it is necessary to monitor spacecraft and satellites in real-time and vital to detect various objects such as space debris.

Due to their small size, weak brightness, and immense distance from the charge-coupled device (CCD) or CMOS sensors, space objects occupy only a few pixels in an optical image and often appear as points and lines [3]. So the detection of space target is the extraction of points or lines according to the different modes of optical telescope. Over the past few decades, many typical algorithms have been proposed and applied in dim and small target detection to address the problem. The extraction of faint objects in space is often a multi-step process that always includes background and noise removal, star removal, and detection of targets at the end.

Space target detection methods can be divided into single-frame and multi-frame algorithms according to different requirements of algorithms. The algorithm based on a single frame directly uses filtering or morphology to extract the target, which can achieve faster detection and real-time tracking. However, when the shape of the target is not apparent, such a method is easy to cause missed detection or false alarm. Multi-frame methods use inter-frame information to judge the authenticity of the target in the image. Compared with single-frame methods, these methods can judge the target more accurately, mainly the target with obscure features or the target blocked by stars. However, multi-frame detection methods rely on inter-frame information. They can only

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng H. Zhu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

process sequential images, the processing time of algorithms is always long, and the calculation consumption is colossal.

Methods based on morphological and motion information were proposed for space target detection. Sun *et al.* [4], [5] presented a morphology method and a Median filterer algorithm to detect space debris. Kouprianov *et al.* [6] proposed a method of fitting star and target trajectory using a point spread function and a logical filtering technique to improve automatic target detection. Cament *et al.* [7] used the Bernoulli filter method to track debris in low Earth orbit. Pradhan *et al.* [8] used the 1.3m Devasthal Fast Optical Telescope (DFOT) to obtain images of space targets, effectively identifying space debris up to 50cm in orbit at 1000km with long exposures. A geometric duality method is proposed for multi-target detection, which is efficient and insensitive to initialization [9]. Zamani *et al.* [10] proposed a method for space target detection using inter-frame matching, which is robust in moving target classification while running in near real-time. Guo et al [11] used the methods of image transformation and energy accumulation to detect space targets, which could achieve specific effects on faint GEO targets.

The imaging characteristics of space objects in images under different observation modes are also different. Researchers have also done a lot of research on streak detection of space images. Virtanen *et al.* [12] proposed a prototype pipeline method, the performance of the pipeline on long streaks is ideal, while the capability for detecting short lines is weaker. Laas-Bourez *et al.* [13] proposed a space target detection algorithm based on mathematical morphology, which combines Top-hat and Hough transform. Levesque *et al.* [14] applied the matched filter to the space target detection, which can effectively detect the long target. WASZCZAK *et al.* [15] used machine learning classifier methods to detect streaks of asteroids in LEO, which can effectively distinguish false alarms from real asteroid targets. Vananti *et al.* [16] completed detecting possible fringe objects by matching the streak with the filter and estimating the image background. Zimmer *et al.* [17] proposed a GPU-accelerated streak detection method, which can effectively improve detection speed and is expected to detect near-earth space objects of magnitude 12-13. Nir *et al.* [18] cross-correlated the image with a straight line template broadened by the system's point spread function to achieve streak detection in space images. These methods can achieve accurate detection of long exposure target, but the detection effect of weak short streak generated by short exposure is often not ideal, and the detection speed is not fast enough.

In recent years, a dramatic increase in the calculation capability of graphics processing unit (GPU) has promoted the development of deep learning. Deep learning has made remarkable achievements in target detection [19], classification [20], and semantic segmentation. FCN [21] is a milestone in semantic segmentation, which realizes the application of deep learning in semantic segmentation through the fully convolutional network. Various semantic segmentation algorithms based on deep learning are proposed and achieved success. U-net [22] is used to solve simple segmentation problems of small samples, such as the segmentation of medical images. It follows the same basic principle as FCN, which adopts the Encoder-Decoder structure and realizes richer information fusion. SegNet [23] does not directly fuse the information of layers of different scales, and it uses pooling with coordinates index to solve the problem of information loss. PSPnet [24]uses spatial pyramid pooling to obtain feature maps with different receptive fields. These maps with different receptive fields are concatenated to complete multi-level semantic feature fusion. An encoder-decoder structure with a large dimensional convolution kernel is proposed in GCN [25].

Some researchers try to apply machine learning and deep learning methods to space object detection. CNN can be trained on light curve observation to classify space objects [26]. Rasit Abay *et al.* [27] applied image pyramid network to space target detection and improved the detection performance of GEO space targets by using subsequent processing. Vittori et al [28] proposed a method of extracting target trajectory based on U-Net, which realized fast image processing, but the target trajectory was usually long. Xi et al [29], [30] applied deep learning to space streak detection and neural network to the pipeline of space target detection.

This paper constructs the space target dataset used to segment the target. The dataset contains training sets and test sets of various SNR. Due to the superficial characteristics of space objects, the image number of the dataset need not be huge, and the network should not be too complex to avoid overfitting. Gaussian noise and shot noise are added to the test set to further test the network's robustness. We propose a Channel and Space Attention U-net (CSAU-Net) for target segmentation according to this principle. The network contains an encoder-decoder structure like many other semantic segmentation networks. The attention module is added to the network to make better use of the low-scale feature map information and enhance the attention to the target. For space targets, a loss function suitable for this study is applied to network training, which is used to solve the imbalance of positive and negative samples in space image segmentation and complex case segmentation of dim and weak targets. Finally, a simple connected component labeling method is applied to extract the target position after the segmentation mask outputs of the network.

The remaining works are organized as follows. The proposed method is described detailed in section 2 and section 3. We describe the structure of the network and training set; simultaneously, we explain in detail how the dataset is created. In section 4, we do some experiments to verify the effectiveness of the algorithm, including the comparison with the current semantic segmentation networks and space target detection algorithm. In section 5, we discuss the performance and boundedness of the algorithm. Finally, we summarize the entire research content in section 6.

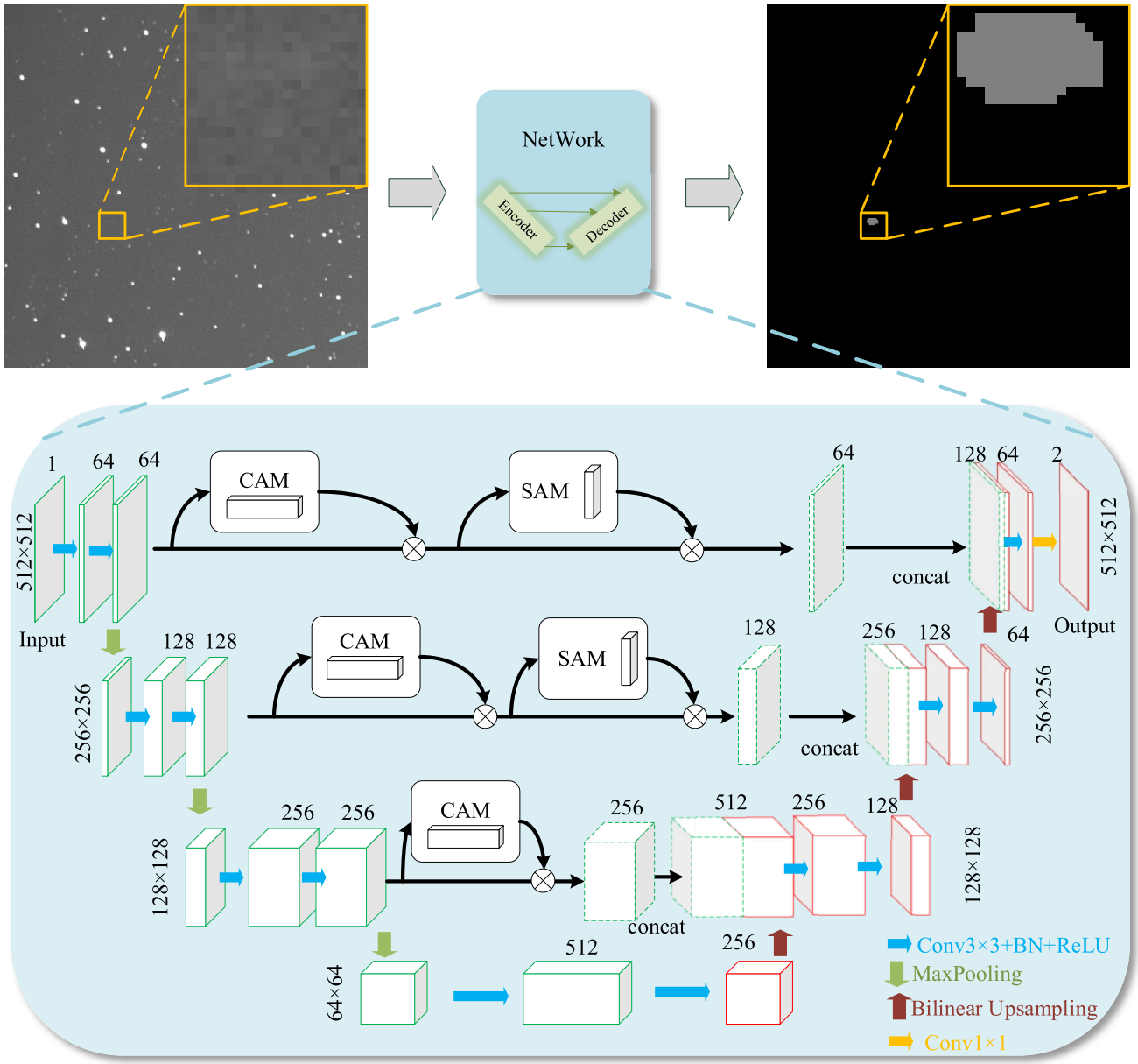


FIGURE 1. The proposed architecture with the segmentation networks. Each cube represents a multi-channel feature map. CAM and SAM are channel attention module and spatial attention module respectively.

II. SEGMENTATION NETWORK

Ground-based telescopes have different observation modes. The method we study is based on the image acquired in the star mode, i.e., the stars appear as points in the picture, while the space objects appear as short streaks after a specific exposure time due to their motion. The image features of space dim target are different from traditional images, the targets that we aim to detect occupy only a few pixels. Not all segmentation networks are suitable for this study because the subsampling structure of convolution and pooling can eliminate small space targets, especially for the backbone network algorithm. Subsequent upsampling layers cannot restore the target information in the image after feature extraction using the backbone network.

The proposed network focuses on segmenting small dim targets on high-resolution images. Therefore, the design of the network should not only avoid the backbone network with too many layers of pooling but also make good use of the feature maps information in front of the architecture effectively. Inspired by U-Net, the proposed network (CSAU-net), like most semantic segmentation networks, is mainly divided into encoder and decoder architecture. The overall structure of the segmentation network is shown in Figure 1. The proposed contracting path is used to reduce the size of feature maps and improve computational efficiency.

Too many pooling layers are not suitable for the performance of the network. Unlike U-Net, our encoder architecture consists of eight convolutional layers and three max-pooing

layers with stride 2 for down-sampling. The convolution kernels of the network are all 3, and each convolution is followed by batch normalization (BN) and a non-linear ReLU layer. Dropout is not utilized in any layers because the weight sharing in convolutional layers and BN layers provide enough regularization. Shallow feature maps are essential to the segmentation because of the small target size.

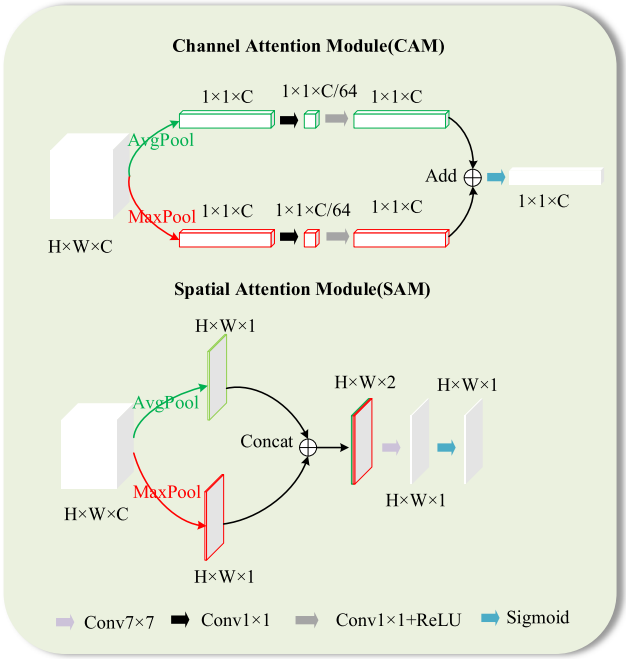


FIGURE 2. Schematic diagram of channel attention and spatial attention module.

We exploit shallow feature maps by introducing attention modules [31] for skip connections. In the channel attention module (CAM), feature vectors with the same channels are pooled in each feature map. Two feature vectors concatenate with each other. The spatial attention module (SAM) is different from it, Corresponding vectors are pooled at each point in the feature map, and finally, a single channel feature map is obtained. Two attention modules provide the feature map's channel and space weighted coefficients, respectively. Channel attention module (CAM) and space attention module (SAM) are shown in the following formula:

$$M_c = \sigma(MLP(AvgPool\ (F)) + MLP(MaxPool(F))) \quad (1)$$

$$M_s = \sigma(f^{7\times7}([AvgPool\ (F)\ ; MaxPool(F)])) \quad (2)$$

where M_c is the weight parameter of the channel attention module, M_s is the output of spatial attention, and F is the input feature map. It is worth noted CAM and SAM are utilized together in the first two skip connections, and only CAM for the third time. In low-size feature maps, spatial attention does not yield better benefits and increases the parameters of the network.

Bilinear interpolation but not the convolutional layer is utilized in the upsampling operation. This method can satisfy the size of the restored feature maps without bringing additional network parameters. The expansive path overlays the feature maps weighted by the attention module with the interpolated image to realize the information fusion. The network's input is a single-channel grayscale image. The output is a 2-dimensional tensor, where the corresponding pixel of each point value is the probability of classification as background or target. The network completes the segmentation of space objects by classifying all image pixels.

III. DATASET AND IMPLEMENTATION DETAILS

A. DATASET CONSTRUCTION

There is no space target dataset with labels available for use at present, and a large amount of labeling of space images will also cause high cost and time consumption. Therefore, we create a dataset for semantic segmentation and space object detection. Unlike traditional pictures with complex scenes and contours, space images can be synthesized and annotated artificially. One hundred target blocks are clipped and added to the image without any target. The requirement for target selection is to select different targets detected from the same detector at other times so that the network can learn more target types as far as possible and avoid overfitting of the network. The clipped target block is a rectangle two pixels larger than the target to prevent the impact of noise and stray light on the target. The method [32], [33] of adding simulation targets to the image cannot perfectly simulate the target, resulting in the overfitting phenomenon of high accuracy of the training set and almost zero accuracies of the test set. The process of adding a target is as follows:

$$F\ (x, y) = f\ (x, y) + \frac{[T\ (x, y) - \min\ (T\ (x, y))]}{k} \quad (3)$$

where $F(x, y)$ is the image after adding the target, $f(x, y)$ is the image without any space target, $T(x, y)$ is the target image block. k is a hyperparameter used to adjust the SNR of the target, and it is uniformly distributed randomly, i.e., $k \sim U(1, 5)$. Another advantage of manually adding space targets to generate a dataset is that various targets with SNR can be added to the image so the network can learn from fainter targets. Traditional annotation needs to pay attention to the unclear target in the image acquisition and the annotation, which is challenging to complete.

We introduce an index of space target weakness to weigh the degree of faintness of the target in the image. The signal to noise ratio (SNR) of the target is defined as follows:

$$\text{SNR} = \frac{E_r - E_B}{\delta_B} \quad (4)$$

where E_r is the mean value of the target region, E_B is the mean value of the background region, and δ_B is the standard deviation of the background region. Generally, the background region is three times the size of the target region. SNR of the target is determined by the intensity and background of the target signal. The larger the SNR is, the more prominent the target is in the image, and the easier it is to be segmented

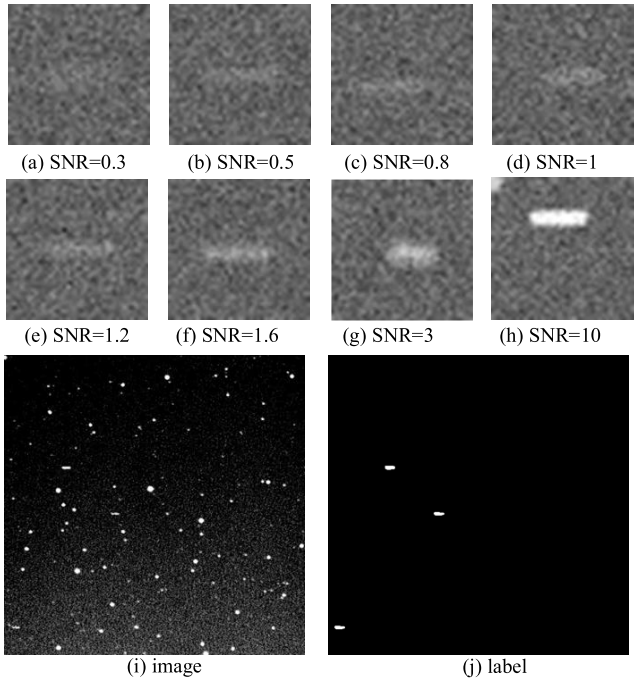


FIGURE 3. (a)-(h) shows targets with different SNR, respectively. (i) the image is processed with image enhancement for better visual effect. (j) segmentation mask.

and detected by the algorithm. Figure 3 shows the images of space targets with different SNR and image pairs for training. In the dataset, 5000 images were used as the training set, 1000 images were used as the verification set, and images with different SNR were used as the test set with 100 images of each type. The target block's position is added to the image and is retained in the test set for verifying the subsequently connected component labeling and centroid extraction.

B. LEARNING DETAILS

Images are normalized before input to the network to enhance the segmentation performance of the network. The network task can be addressed as a binary-image classification problem. Different loss functions have a significant influence on the training effect of the network. Focal Loss [34] is an improvement of cross-entropy loss as one of the loss functions. On the other hand, stars and targets differ in shape and grayscale. Dice coefficient [35] is an index to evaluate image overlap as another loss function in this paper. Focal Loss and Dice loss are defined as follows:

$$L_{Focal\_loss} = -\left(\alpha\left(1-\hat{y}\right)^{\gamma} y \log\left(\hat{y}\right) + (1-\alpha)\hat{y}^{\gamma}(1-y)\log(1-\hat{y})\right) \quad (5)$$

$$\text{Dice} = \frac{2\sum_{i=1}^{2} y_i \cdot \hat{y}_i}{\sum_{i=1}^{2} y_i^2 + \sum_{i=1}^{2} \hat{y}_i^2}, L_{Dice} = 1 - \text{Dice} \quad (6)$$

where y and $\hat{y}$ are the ground truth and predicted value of the network, respectively. α and γ are hyperparameters of Focal Loss for different categories and different classification difficulties. There is a severe imbalance of positive and negative samples in the space target image. The background occupies most of the image, while the targets only take up dozens of pixels. To assess the distribution of targets and background, the calculation of α is as follows:

$$\alpha = \frac{1}{N}\sum_{i=1}^{N} \frac{\log(p + \frac{N_t^{(i)}}{N_t^{(i)}+N_b^{(i)}})}{\log(p + \frac{N_b^{(i)}}{N_t^{(i)}+N_b^{(i)}})} \quad (7)$$

where p is the hyperparameter used to adjust the value of α, and we set it as 1.10; N_t and N_b are the number of pixels occupied by the target and background in the mask; N is the total number of images used in the training set. The parameter of this dataset is 0.13 after calculation. The value of γ is 2, the same as Lin [34]. The final loss function is shown as follows:

$$L_{total} = k_1 L_{Focal\_loss} + k_2 L_{Dice} \quad (8)$$

where k_1 and k_2 are the hyperparameters used to adjust the two loss function, which we set to 0.8 and 0.2. Focal Loss is set larger because it allows the network to focus more on classifying targets and the segmentation of challenging targets at the pixel level.

There are various choices of optimizers. After experimental comparison, Adam can get training results faster, but the final convergence effect is often not as good as the SGD optimizer. To achieve a faster training effect and better convergence result, we use the training strategy of Adam first and then SGD. According to Keskar [36], switching between two optimizers as early as possible can get better results. We choose to train the network for 150 epochs with SGD after 50 epochs of training with Adam. The learning rate and decay rate of the two optimizers are all 1×10^{-3} and 2×10^{-6}.

C. COORDINATE CALCULATION

The binary mask is obtained after inference of the convolutional neural network. However, we often want the result to be the specific location of the target. A simple connected component labeling method is used to extract the location of the space target.

Firstly, the mask image is scanned line by line. A sequence of continuous white pixels in each line is called run, which records its starting and ending points. Give a new label to a run in all rows except the first if it has no overlap with any run in the previous row; If it has overlapped with only one run in the previous row, it is assigned the label of that run in the previous row; If it overlaps with more than two runs in the previous row, the current run is assigned a minimum label of the contiguous run, and the marks of these runs in the previous row are written into equivalent pairs, indicating that they belong to the same category. To convert an equivalent pair to an equivalent sequence, each sequence needs to be given the same label because they are all equivalent. Each equivalent sequence is assigned a label starting at the first row. We iterate over the label of the first run, find the equivalent sequences, give them a new label and fill the label of each group into

the label image. We can quickly get the pixel position of each target after completing the mark of the connected component. The median coordinate of each component is taken as the specific coordinate of the target.

It was considered to add a predict head for regression coordinates to the network's back end. When the segmentation mask is not ideal, such as incorrectly dividing noise into targets, the extra network can get the correct segmentation and achieve more accurate results. However, the results of the semantic segmentation network can get accurate results. The additional network structure will not only not improve the detection accuracy but also lead to an increase in network parameters. The above simple connected component labeling method is sufficient to obtain the exact coordinates of the target, and the additional network structure is unnecessary.

IV. EXPERIMENTS

We first introduce performance metrics in this section and then present our experiments in detail. We discuss the influence of data pre-processing on the algorithm and verify the robustness by enhancing the dataset. Several classical semantic segmentation algorithms are compared with CSAU-Net under the same training strategy to compare the segmentation effectiveness of the network. Finally, CSAU-Net is tested on real images and compared with traditional space target detection methods, proving the algorithm's feasibility. Our experiments were conducted on a computer with 16-GB random access memory, Intel Core i7 8700K, 3.6GHz processor, and Nvidia 2080ti GPU. The network architecture was implemented in Pytorch 1.7.

A. PERFORMANCE METRICS

In this paper, to verify the segmentation performance of the convolutional neural network and the detection performance of the algorithm, the following evaluation criteria are introduced:

1) DICE COEFFICIENT AND MEAN INTERSECTION OVER UNION (MIOU)

The calculation method of the Dice coefficient is shown in (6). Dice will calculate the product of the corresponding pixels of ground truth (GT) and the prediction mask. The correct classification of non-zero pixels in GT, namely target pixels, is the focus of the Dice coefficient, but pixels incorrectly classified as targets in the prediction mask are ignored. Therefore, we introduce another evaluation standard, MIoU, whose expression is:

$$\mathrm{MIoU} = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}+\sum_{j=0}^{k}p_{ji}-p_{ii}} \tag{9}$$

where k is the number of categories (including empty classes), p_{ii}, p_{ij}, p_{ji} represent correctly classified pixels, false positive and false negative, respectively. MIoU calculates the classification effect of each category by calculating the confusion matrix, which evaluates the segmentation performance of various categories globally.

2) DETECTION RATE AND FALSE ALARM

MIoU and Dice coefficients jointly evaluate the segmentation effect of semantic segmentation networks on images. We only use them to compare various convolutional neural networks in this paper. The target coordinates obtained through the connected component labeling algorithm are the final results we hope to get. Only the central coordinate as the location of the target is needed since the space target is tiny in the image. When the Euclidean distance between the coordinates obtained by the algorithm and GT is less than 10, the target is considered to have been detected successfully. To evaluate the overall detection performance of the algorithm, detection rate (P_d) and false alarm (P_f) are introduced:

$$P_d = \frac{N_d}{N_{all}}, \quad P_f = \frac{N_f}{N_{all}+N_f} \tag{10}$$

where N_d represents the number of correctly identified targets, N_{all} represents total targets, and N_f represents the number of stars or noise points wrongly identified as targets.

B. EXPERIMENT ON TRAINING METHODS

Convolutional neural networks often have different fitting effects on the dataset with different data distributions, and different training strategies have a specific influence on the final convergence results of the network. The network is evaluated according to the input data and training methods and divided into the following four groups of experiments:

1. Input images with three kinds of resolutions (256 × 256, 512 × 512.1024 × 1024)
2. Two loss functions for training (Cross-Entropy and Focal Loss + Dice loss)
3. Data pre-processing (Normalization and Standardization)
4. The impact of the optimizer (Adam, SGD, RMSprop, Adam+SGD).

1) DIFFERENT RESOLUTION OF THE INPUT IMAGE

Increasing the image's resolution can improve the texture of the target and other areas in the image so that the convolutional neural network can learn more features. Reducing image resolution can minimize computation costs and speed up network reference and image segmentation. Images with three resolutions were set as tests in this experiment. Images with different resolutions will be uniformly resized to the same size as original images after the network's output to ensure the fairness of subsequent detection.

As shown in figure 4 (a), the detection accuracy of halved resolution is significantly reduced, especially for targets with SNR lower than 0.8. The false alarm rate with half-resolution is also more prominent than the original resolution. Although the detection accuracy of the image after linear interpolation is higher than or equal to the original image when the target SNR is higher than 1, its performance is not ideal when the

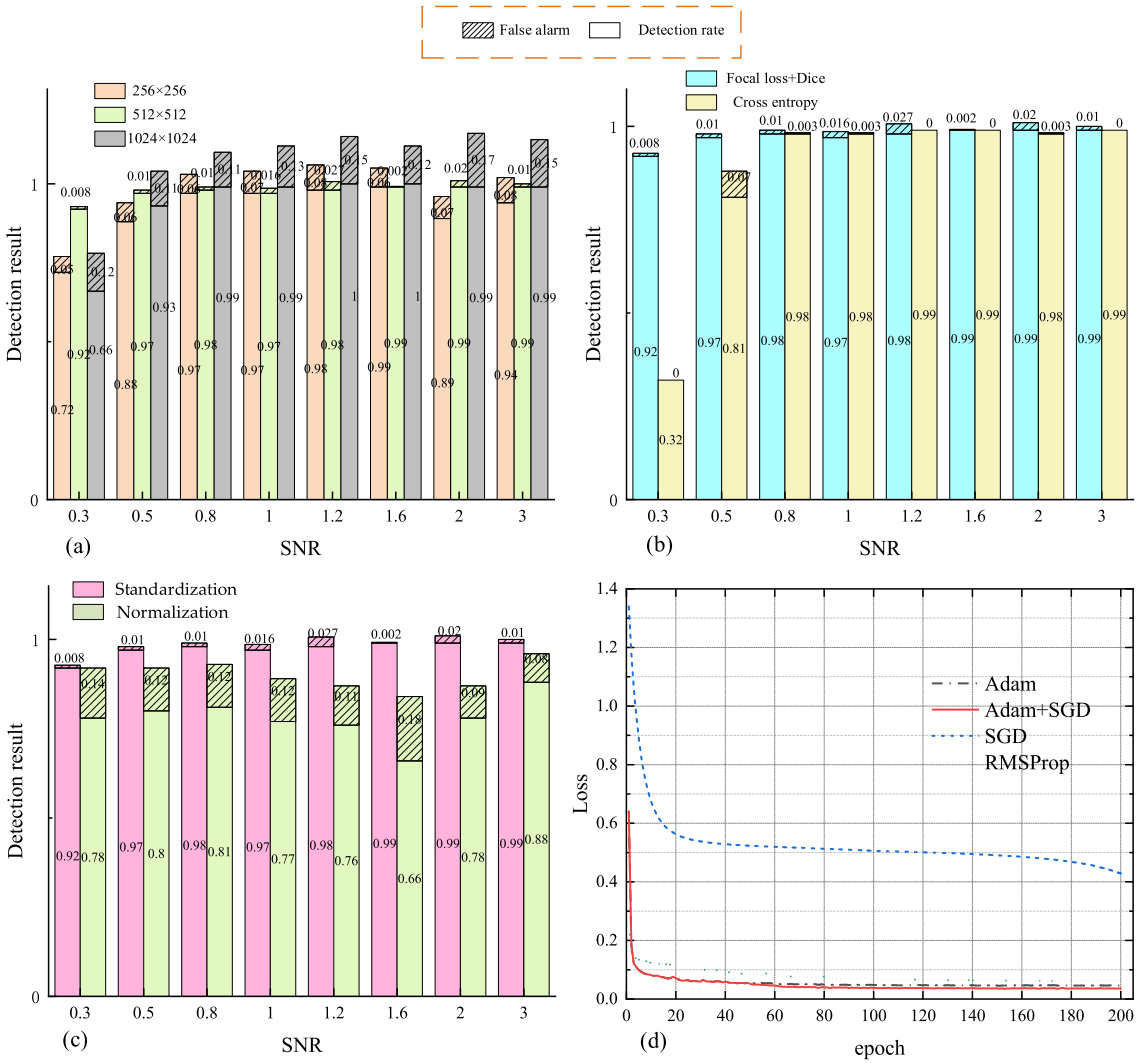


FIGURE 4. Results of different preprocessing or training methods. (a) the detection results of images with different resolutions input by the network; (b) the comparison between Focal loss+Dice and CE as the loss function; (c) the comparison between the network performance after normalization and standardization of input; (d) loss-epoch curve of the four optimizer training strategies.

TABLE 1. Inference time of different resolution.

resolution	256×256	512×512	1024×1024
Inference time	2.7ms	2.8ms	3.4ms

SNR is extremely low. The detection accuracy is only 66% when the SNR is 0.3. At the same time, its false alarm is higher than the other two at each SNR. The reason for this phenomenon may be that the image interpolation enhances the characteristics of noise while enhancing the size of the target, leading to the network wrongly classifying more noise pixels as targets.

As shown in Table 1, the average network inference time after halving the image is only 0.1 ms less than the original resolution, and the inference time after image interpolation is 3.4 ms. The inference time of three resolutions can meet the requirements of real-time segmentation and detection, but considering various SNR targets, the segmentation and detection performance of original image resolution is the best choice.

2) DIFFERENT LOSS FUNCTION

Cross-Entropy(CE) loss has an excellent performance in various algorithms as a classical loss function of semantic segmentation [22]. Figure 4 (b) compares CE loss as loss function and Focal Loss + Dice as loss function. The network using CE loss as loss function has the same segmentation effect on obvious targets as our method, but the segmentation effect on targets with SNR lower than 1 is poor. When the SNR of the target is 0.3, the detection rate of the algorithm is only 32%, and when the SNR is 0.8, the detection rate of the

algorithm is only 81%, far lower than the method used in this paper.

CE loss does not focus on hard-to-segment cases in the network's training, and there are few target pixels in space images, resulting in a severe imbalance of positive and negative samples. CE loss in the optimization of the network will make the network pay more attention to the pixels that are easy to classify, resulting in the network not generating many false alarms during pixel classification in most cases.

3) DATA PRE-PROCESSING COMPARISON

Normalization and standardization are both necessary means of data pre-processing in machine learning, as shown in the following formula:

$$x_{\text{nor}} = \frac{x}{\max(x)}, \quad x_{\text{str}} = \frac{x - \text{mean}(x)}{\max(x)} \tag{11}$$

Figure 4 (c) shows the detection results of the two pre-processing methods. The normalized dataset is still unevenly distributed, which causes great difficulty to the convergence of the network, and it is difficult to achieve the ideal effect in the segmentation of space targets. In the image of various target SNR, the standardized dataset is more conducive to network fitting of data. The segmentation image detection accuracy is higher; meanwhile, the false alarm rate is lower.

4) COMPARISON OF DIFFERENT OPTIMIZERS

The network is trained using RMSprop, Adam, and stochastic gradient descent (SGD) alone as optimizers and a combination of the two. Under the condition that other factors are equal, the number of training rounds is 200. Figure 4 (d) shows the loss-epoch curves for the four optimizer strategies. The network trained by Adam and RMSProp converges quickly, but the convergence effect is not as good as the method in this paper. The convergence rate of SGD is slow, so the verification result of the final network is poorer than the other two methods. Adam is used to optimize the parameters to roughly achieve the fastest convergence speed. Finally, SGD is used to fine-tune the parameters to obtain the optimal solution satisfying the target segmentation. Compared with Adam or SGD alone, combining the two optimizers makes the network converge quickly and achieves high accuracy.

Different data pre-processing methods and training strategies significantly impact network training. Although the training method in this paper does not achieve the best segmentation and detection results in some cases, the algorithm in this paper can keep a low false alarm rate while ensuring a high detection rate at each SNR.

C. ABLATION EXPERIMENTS

An ablation experiment is performed to verify the effect of the attention mechanism on network segmentation results. We compare six different network structures: (1) original U-Net; (2) Remove U-Net of low-resolution feature maps (modified U-Net); (3) SAM structure is added separately to the modified U-Net; (4) CAM structure is added separately in the modified U-Net; (5) CAM structure and SAM structure (CSAU-NET) are added to U-NET; (6) CAM structure and SAM structure (CSAU-NET) are added to the modified U-NET. The training details are the same as the training set to ensure the fairness of the experiment. We only select the image with a SNR of 1 as the test set because the target can well represent the dark and weak target in practical detection.

TABLE 2. Detection results of networks with different attention modules.

Description	P_d (%)	P_f (%)
U-Net	90.5	7.6
Modified U-Net	88.6	12.3
Modified U-Net+SAM	96.5	5.8
Modified U-Net+CAM	98.6	3.6
U-Net+CAM+SAM	99.2	3.6
Modified U-Net+CAM+SAM	99.2	2.1

It can be seen from the experiment that the segmentation effect of the modified U-Net network is reduced due to the removal of part of the network structure. After adding the attention module, the network performance is improved to different degrees. The network segmentation effect by adding SAM attention module is better than that by adding CAM module. The characteristics of space targets are weakened in pooling, leading to a more significant effect on the spatial weighting of different channels. After adding two attention modules, the performance of the network is greatly improved. Among hundreds of targets, only one target failed to be segmented successfully.

D. ROBUSTNESS ANALYSIS

Space detectors are susceptible to various influences when acquiring space images, like background noise, Photon count errors, etc. To explore the robustness of the proposed method, the network is tested by different images, including images with additional Gaussian noise and shot noise.

Two different training sets are used for network training. The training set without any noise is used to train the network model firstly, and the model is used to segment the images with different noises. Then we randomly add various noises to the training set. Only the target with an SNR of 1 is taken as the test set in this experiment to eliminate the influence of different targets on the results. The peak signal-to-noise ratio (PSNR) is used to evaluate the degradation degree of the image after adding noise, as shown in the following formula:

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)] \tag{12}$$

$$\text{PSNR} = 10 \cdot \log\left(\frac{255^2}{\text{MSE}}\right) \tag{13}$$

where I and K are the images before and after noise pollution, respectively, m and n are the image size. PSNR is different from the previously defined SNR. SNR expresses the

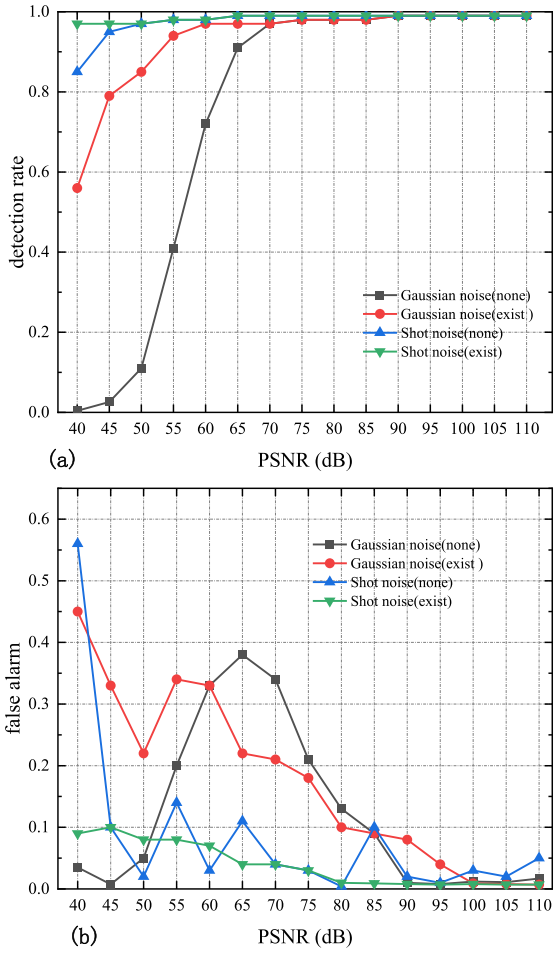


FIGURE 5. Detection result of the algorithm on images with different noises. (a) and (b) are respectively the detection rate and false alarm under the influence of different noises. "none" and "exist" represent whether the training set contains noise.

significance degree of spatial objects, while PSNR describes the comparison of images before and after noise pollution. The larger the value is, the lower the degradation degree of the image is.

Figure 5 shows the image detection results of the network model trained by two training sets. With the increase of various noises, the detection rate of the network gradually decreases, and the false alarm gradually increases. It is difficult for human eyes to distinguish space targets in degraded images with PSNR less than 70 after adding Gaussian noise. It can be seen from Figure 5 (a) that the network has a significantly better generalization effect than Gaussian noise on images added with shot noise. The detection rate of images added with shot noise can reach above 0.9 when the PSNR of the image is greater than 55dB. After adding noise to the training set, the network can be well adapted to both kinds of noise, especially to the image polluted by shot noise, and the network can stably segment the target. Figure 5 (b) shows the false alarm of different images detected by the proposed method. The reduction of noise in the image will lead to the reduction of false alarm. However, the network tends to classify all pixels of the image as background in the image with severe Gaussian noise pollution, so the detection rate is not ideal, and the false alarm is low. With the decrease of Gaussian noise, the network will gradually misclassify some pixels as targets, increasing the false alarm rate. Finally, the network can correctly classify the pixels with the gradual disappearance of noise. After adding noise to the training set, the network segmentation effect improves, and a more miniature false alarm is realized.

E. COMPARISON WITH DIFFERENT SEGMENTATION NETWORKS

Many deep learning network models for semantic segmentation, but not all networks are suitable for space object segmentation. In most cases, the more complex the network is, the better the result in traditional target detection and semantic segmentation tasks. However, the effect of complex backbone-based semantic segmentation methods such as Deeplab v3+ and Segnet are often poor in our study. In the study of space target segmentation using backbone like resnet-101 or vgg-16, the network tended to classify all image pixels as the background. It could not complete the segmentation task through experiments. There are two main reasons for this phenomenon.

1) DOWNSAMPLING LAYERS

The images to be processed in the network are high-orbit space target images obtained by ground-based telescopes, so the target pixels in the picture are few, usually only a few to dozens of pixels. The complex network can extract complicated image features, but the space targets do not have complex features or textures. The backbone structure can not improve the segmentation accuracy and reduce the number of pixels of the space target with too many pooling layers, resulting in the target not being effectively segmented.

2) DATASET

Complex networks often need a large amount of data for training to enhance generalization and reduce the problem of network overfitting. The image features of this study are simple, and most observation situations can be covered without a large number of datasets, so the possibility of overfitting is greater.

FCN [21], U-Net [22], and ESP-Net [37] are chosen for network comparison. The training strategy of U-Net and ESP-Net is the same as that of this algorithm. FCN adopted in this paper takes the last four layers of Resnet-34 as feature extraction, so the weight trained on ImageNet is used as the initialization parameter for network parameters. It is worth noting that the feature extraction layer of FCN does not completely use the backbone to achieve good segmentation results in this dataset. To verify the segmentation effectiveness of the proposed method, we test images with different SNRs. After training the same dataset, various networks are tested using the common test set. The number of space targets in each image varies from 0 to 3 to realize that the simulation

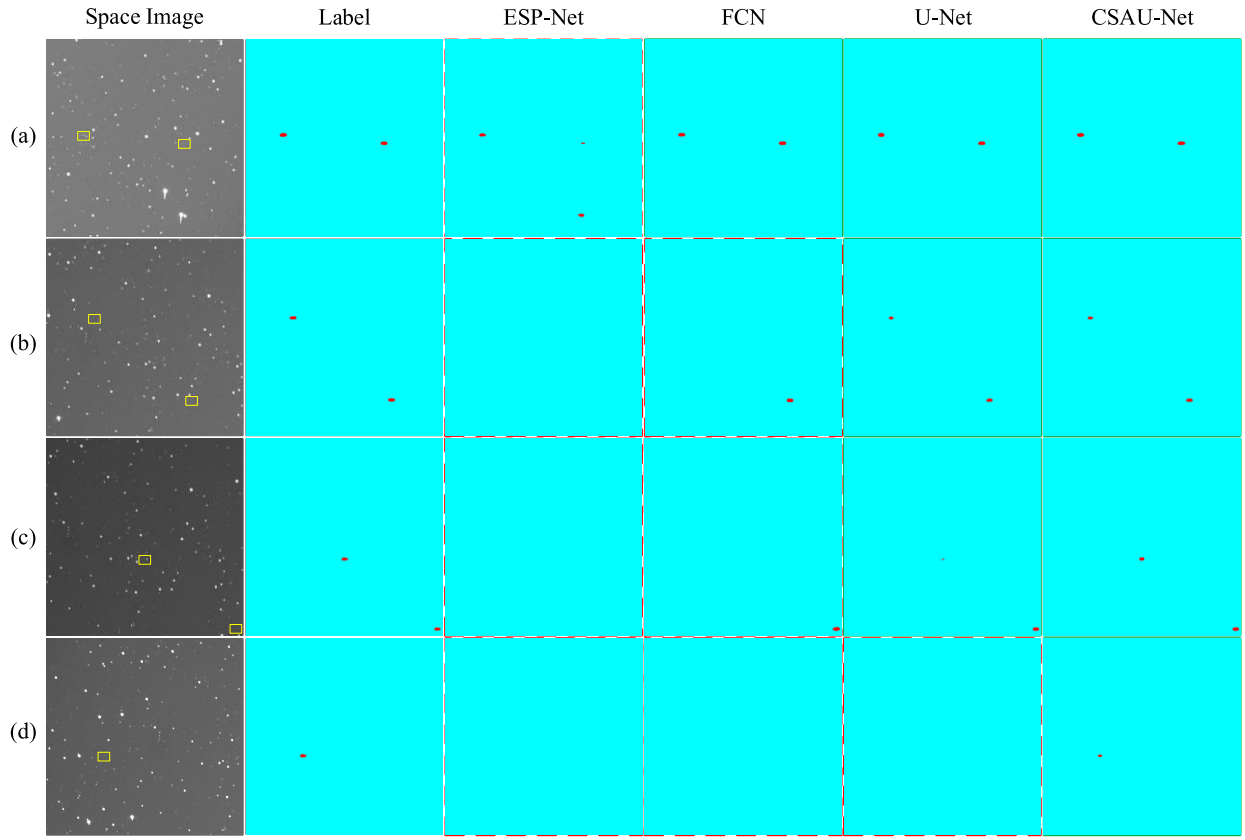


FIGURE 6. The segmentation result of different networks. (a)–(d) are four space images with different SNR targets. The yellow rectangle represents space target in image, the red part in the segmentation mask is the target segmented by the network.

dataset is closer to the real data, similar to the training sets' distribution. Figure 6 shows the segmentation result of each network. All the networks have perfect segmentation effectiveness on the target with obvious high SNR. For example, for the two targets in the group (a) experiment, most networks achieved segmentation perfectly, except that ESP-Net mistakenly segments the part of the background as the target. With the reduction of target SNR, the segmentation task becomes more difficult. ESP-Net and FCN perform poorly in the face of extremely low SNR. In the segmentation experiment of the group (b) and (c), ESP-Net could not segment the target normally, and the network classifies all pixels as background. FCN also can only segment one target in the group (c). Similarly, segmentation of U-Net is also not ideal, and only CSAU-Net successfully segments targets. In group (d) experiment, only CSAU-Net complete the segmentation of space targets. When the target SNR is lower than 0.5, only CSAU-Net can segment the target stably. U-Net can achieve a good effect as a representative of the segmentation network. Our modification on U-Net makes it easier to segment weak space targets.

When the segmentation effect is poor, and the networks classify most of the image area as targets, many false alarms will be generated. In this case, the detection rate and false alarm rate have little significance as evaluation criteria, and we adopt Dice and MIoU as evaluation criteria. The specific segmentation data is shown in figure 7. To express the effect of the algorithm more clearly, we annotate the result of CSAU-Net. When the SNR of the target is more prominent than 0.5, the two evaluation indexes of CSAU-Net achieve a high effect, and networks perform accurate segmentation of the target. Compared with other semantic segmentation methods, the proposed method has more advantages in target segmentation when the SNR is less than 1.6. When the SNR of the target is high and the target is obvious, multiple segmentation methods can segment the target, but the result of CSAU-Net is the most accurate.

The inference time of the algorithm is also an essential evaluation of the algorithm. Inference time of ESP-Net and FCN is 5.6 ms and 3.5 ms for the image with 512×512 resolution, respectively. CSAU-Net adds a small number of parameters compared to traditional U-Net. Due to the addition of the attention module, the inference time of the network is 0.4ms longer on the device we use, but this is tolerable relative to the improvement in segmentation accuracy. The exposure time of the detector observing the space target is usually several seconds. The proposed algorithm can ultimately realize the online segmentation of the target.

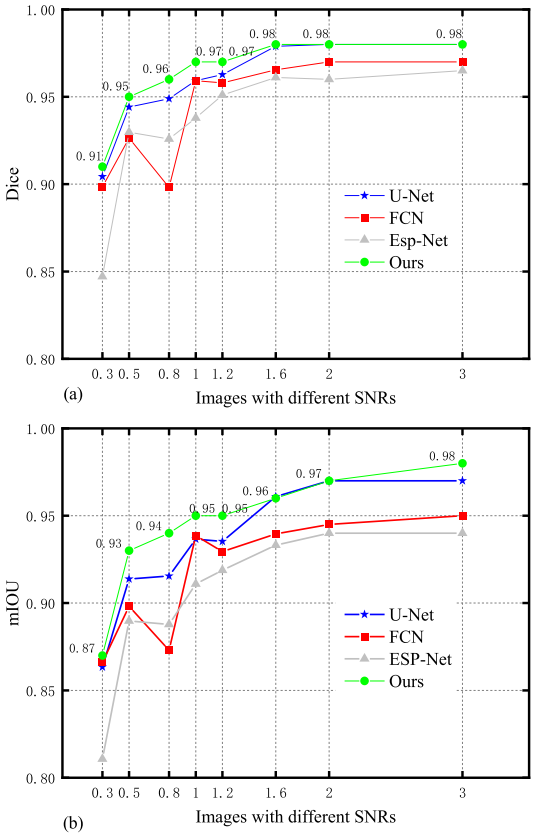


FIGURE 7. Segmentation results of images with different SNRs by different networks. (a) and (b) are Dice and mIOU of the proposed method's results respectively.

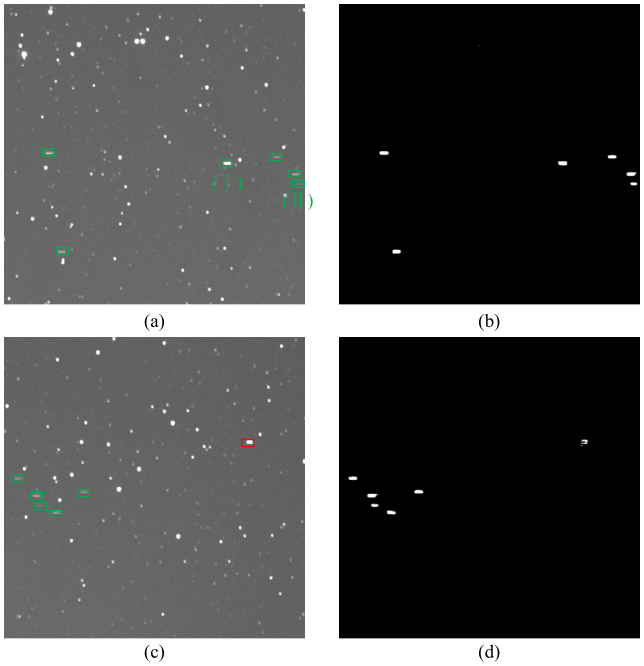


FIGURE 8. Network segmentation of images containing real space objects. (a) and (c) are images containing space targets, (b) and (d) are segmentation results. Green rectangles contain the real targets, and red rectangles contain stars misclassified as targets in (a) and (c).

F. COMPARISON WITH OTHER TARGET DETECTION METHODS

The images used in this experiment are space images obtained by an optical telescope which has a $10° \times 10°$ field of view with a 5 s exposure time. The images are 16-bit gray images with a resolution of 2048×2048. Images with different SNRs (0-10) are selected as the test, and the data are cropped to the image with a resolution of 512×512, with the number of space targets in each image ranging from 0 to 10. Twenty sequences of imagescontaining 100 images were tested to compare various algorithms.

We test images containing natural space objects and study the detection performance of the algorithm on real images. The segmentation effect of the network on the actual image is shown in Figure 8. The network trained only with the training set without noise can effectively segment space targets with different degrees of faintness. The image contains space targets with different SNRs. The SNR of the target (I) is 15, while that of the target (II) is only 0.5 In Figure. 8 (c). The traditional thresholding-based segmentation method often divides the weak target into the background while the apparent target is segmented, which reduces the detection rate. Our approach can achieve stable segmentation of dark and dim objects near bright objects.

The network classifies pixels based on the morphological characteristics of the image. Hence, the network mistakenly divides the stars into targets when two stars are linked together and look like targets. It is the disadvantage of the single-frame segmentation method. Using the multi-frame association method can remove false alarms.

Some space target detection methods in recent years are selected for comparison to verify the comparison between the proposed method and traditional methods. The chosen methods include multi-frame target detection method and single-frame target detection method. Multistage hypothesis testing (MHT) [38] is an algorithm to detect small moving targets with unknown prior knowledge. DPSWM [33] is an improved dynamic programming sliding window method to detect space targets, achieving higher detection accuracy. Among them, FLCR [30], Vittori [28], and FGBNN [29] are based on deep learning methods.

The multi-frame detection method can further judge whether there is a target in the image after processing the image with the information of multiple frames. In contrast, the single-frame method only detects the space target according to the knowledge of a single frame. Table 3 shows the detection results of each algorithm on the test data set. Compared with the single-frame method, multi-framemethods can use more information, thus achieving a higher detection rate than the standard single-frame method. However, multi-frame methods need to simultaneously judge the relationship between frames, consuming a lot of time and computing resources. The method proposed in this paper can guarantee real-time detection and achieve a better detection effect.

V. DISCUSSION

The method proposed in this paper can train a robust network model with fewer data. CSAU-Net can more accurately

TABLE 3. Detection results of real images by different methods.

Methods		P_d (%)	P_f (%)
Multi-frame	MHT	88.5	14.7
	TS[9]	**99.2**	3.5
	DPSWM	98.3	5.3
	FLCR	98.5	2.4
Single-frame	Vittori	92.8	4.6
	FGBNN	88.6	1.9
	Vananti[16]	94.5	6.5
	Nir[18]	93.6	5.7
	Ours	98.5	**1.6**

segment targets with various SNR than other deep learning methods. Unlike traditional single-frame space target detection methods, the proposed method can achieve high-precision detection even when the target is imaged as a short target (length $<$ 20 pixels) in the image. Other single-frame detection methods often require a long target (length $>$ 50 pixels). It is also why the traditional single frame detection method has a low detection rate in comparison experiments. As the method in this paper detects the target based on a single frame image, it can achieve a faster detection speed. However, when the target features are not apparent due to the insufficient exposure time of the detector, the algorithm cannot be applied at this time. Convolution neural network needs feature extraction of characteristics. Sensors of star mode can obtain the target of the image appearing as strips, the network can effectively make use of target characteristics, but the network is challenging to get the target characteristics of the faint target in the staring mode of the detector so that segmentation ability will be greatly reduced.

VI. CONCLUSION

In this study, we propose a single-frame space target detection method. Firstly, we present an improved encoder-decoder convolutional neural network to complete the feature extraction and segmentation of space images. The network adds an attention module to the fusion of different feature map information to better utilize the original feature layer. The network architecture achieves end-to-end segmentation of space targets without multiple steps of traditional methods. At the same time, a small space image dataset is constructed for segmentation and detection to complete the algorithm's training. The dataset contains images of space objects with various SNRs, which improves the network's generalization. Finally, a connected component labeling method for centroid extraction is applied, which realizes the extraction of the specific location of the target from the mask after network inference.

We complete multiple sets of experiments to verify the performance of the algorithm. First, we compare different dataset processing methods and network training strategies and choose the most suitable loss function and optimizer for this study. Then we add two different noises to the test set to simulate the image degradation that the detector may cause when acquiring images and verify the algorithm's robustness. Finally, we compare different methods. The proposed method can segment space targets better than other semantic segmentation networks and quickly detect targets with multi-frame detection accuracy.

REFERENCES

[1] G. Tommei, A. Milani, and A. Rossi, "Orbit determination of space debris: Admissible regions," *Celestial Mech. Dyn. Astron.*, vol. 97, no. 4, pp. 289–304, Feb. 2007.

[2] H. Wirnsberger, O. Baur, and G. Kirchner, "Space debris orbit prediction errors using bi-static laser observations. Case study: Envisat," *Adv. Space Res.*, vol. 55, no. 11, pp. 2607–2615, Jun. 2015.

[3] B. Sease, B. Flewelling, and J. Black, "Automatic streak endpoint localization from the cornerness metric," *Acta Astronautica.*, vol. 134, pp. 345–354, May 2017.

[4] R.-Y. Sun, J.-W. Zhan, C.-X. Zhao, and X.-X. Zhang, "Algorithms and applications for detecting faint space debris in GEO," *Acta Astronautica*, vol. 110, pp. 9–17, May/Jun. 2015.

[5] R.-Y. Sun and C.-Y. Zhao, "A new source extraction algorithm for optical space debris observation," *Res. Astron. Astrophys.*, vol. 13, no. 5, pp. 604–614, May 2013.

[6] V. Kouprianov, "Distinguishing features of CCD astrometry of faint GEO objects," *Adv. Space Res.*, vol. 41, no. 7, pp. 1029–1038, Jan. 2008.

[7] L. Cament, M. Adams, and P. Barrios, "Space debris tracking with the Poisson labeled multi-Bernoulli filter," *Sensors*, vol. 21, no. 11, p. 3684, May 2021.

[8] B. Pradhan, P. Hickson, and J. Surdej, "Serendipitous detection and size estimation of space debris using a survey zenith-pointing telescope," *Acta Astronautica*, vol. 164, pp. 77–83, Nov. 2019.

[9] D. Liu, B. Chen, T.-J. Chin, and M. G. Rutten, "Topological sweep for multi-target detection of geostationary space objects," *IEEE Trans. Signal Process.*, vol. 68, pp. 5166–5177, 2020.

[10] Y. Zamani, J. Amert, T. Bryan, and N. Nategh, "A robust vision-based algorithm for detecting and classifying small orbital debris using on-board optical cameras," in *Proc. Adv Maui Opt. Space Surveill. Technol. Conf.*, no. M19-7620, 2019.

[11] L. Guo, W. Zhang, Z. Wang, X. Sun, and Y. Shang, "Weak GEO satellite target detection based on image transformation and energy accumulation," in *Proc. 4th Int. Conf. Image Graph. Process.*, Jan. 2021, pp. 52–58.

[12] J. Virtanen, J. Poikonen, T. Säntti, T. Komulainen, J. Torppa, M. Granvik, K. Muinonen, H. Pentikäinen, J. Martikainen, J. Näränen, J. Lehti, and T. Flohrer, "Streak detection and analysis pipeline for space-debris optical images," *Adv. Space Res.*, vol. 57, pp. 1607–1623, Apr. 2016.

[13] M. Laas-Bourez, G. Blanchet, M. Boër, E. Ducrotté, and A. Klotz, "A new algorithm for optical observations of space debris with the TAROT telescopes," *Adv. Space Res.*, vol. 44, no. 11, pp. 1270–1278, Dec. 2009.

[14] M. P. Levesque, "Image processing technique for automatic detection of satellite streaks," DRDC Valcartier, Tech. Rep. TR 2005-386, 2007.

[15] A. Waszczak, T. A. Prince, R. Laher, F. Masci, B. Bue, U. Rebbapragada, T. Barlow, J. Surace, G. Helou, and S. Kulkarni, "Small near-earth asteroids in the palomar transient factory survey: A real-time streak-detection system," *Publications Astronomical Soc. Pacific*, vol. 129, no. 973, Mar. 2017, Art. no. 034402.

[16] A. Vananti, K. Schild, and T. Schildknecht, "Improved detection of faint streaks based on a streak-like spatial filter," *Adv. Space Res.*, vol. 65, no. 1, pp. 364–378, Jan. 2020.

[17] P. C. Zimmer, M. R. Ackerman, and J. T. McGraw, "GPU-accelerated faint streak detection foruncued surveillance of LEO," in *Proc. AMOS Tech. Conf.*, 203. [Online]. Available: http://www.amostech.com/TechnicalPapers/2013.cfm

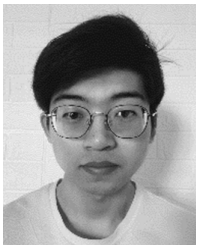
[18] G. Nir, B. Zackay, and E. O. Ofek, "Optimal and efficient streak detection in astronomical images," *Astronomical J.*, vol. 156, no. 5, p. 229, Oct. 2018.

[19] S. Ren, K. He, R. Girshick, and S. Jian, "Faster R-CNN: Towards Real-Time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, vol. 28.

[20] T. Technicolor, S. Related, T. Technicolor, and S. Related, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012.

[21] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[22] O. Ronneberger, P. Fischer, and T. J. S. I. P. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Cham, Switzerland: Springer, 2015, pp. 234–241.
[23] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[24] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2881–2890.
[25] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters—Improve semantic segmentation by global convolutional network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4353–4361.
[26] R. Furfaro, T. Campbell, R. Linares, and V. Reddy, "Space debris identification and characterization via deep meta-learning," in *Proc. 1st Int. Orbital Debris Conf.*, 2019. [Online]. Available: *https://www.hou.usra.edu/meetings/orbitaldebris2019/orbital2019paper/pdf/6123.pdf*
[27] R. Abay and K. Gupta, "GEO-FPN: A convolutional neural network for detecting GEO and near-GEO space objects from optical images," in *Proc. 8th Eur. Conf. Space Debris (Virtual)*, Darmstadt, Germany, Apr. 2021, pp. 20–23.
[28] A. De Vittori, R. Cipollone, P. Di Lizia, and M. Massari, "Real-time space object tracklet extraction from telescope survey images with machine learning," *Astrodynamics*, vol. 4, pp. 1–14, Apr. 2022.
[29] Y. Xiang, J. Xi, M. Cong, Y. Yang, C. Ren, and L. Han, "Space debris detection with fast grid-based learning," in *Proc. IEEE 3rd Int. Conf. Safe Prod. Informatization (IICSPI)*, Nov. 2020, pp. 205–209.
[30] J. Xi, Y. Xiang, O. K. Ersoy, M. Cong, X. Wei, and J. Gu, "Space debris detection using feature learning of candidate regions in optical image sequences," *IEEE Access*, vol. 8, pp. 150864–150877, 2020.
[31] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19.
[32] J. Xi, D. Wen, O. K. Ersoy, H. Yi, D. Yao, Z. Song, and S. Xi, "Space debris detection in optical image sequences," *Appl. Opt.*, vol. 55, no. 28, pp. 7929–7940, 2016.
[33] M. Li, C. Yan, C. Hu, C. Liu, and L. Xu, "Space target detection in complicated situations for wide-field surveillance," *IEEE Access*, vol. 7, pp. 123658–123670, 2019.
[34] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 99, pp. 2999–3007, Jul. 2017.
[35] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.
[36] N. S. Keskar and R. Socher, "Improving generalization performance by switching from Adam to SGD," 2017, *arXiv:1712.07628*.
[37] S. Mehta, M. Rastegari, A. Caspi, L. Shapiro, and H. J. S. Hajishirzi, "ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Cham, Switzerland, 2018, pp. 552–568.
[38] S. D. Blostein and T. S. Huang, "Detecting small, moving objects in image sequences using sequential hypothesis testing," *IEEE Trans. Signal Process.*, vol. 39, no. 7, pp. 1611–1629, Jul. 1991.



XIANGJI GUO received the B.S. degree in mechanical engineering from Jilin University (JLU), Changchun, China, in 2018. He is currently pursuing the Ph.D. degree in mechanical manufacturing and automation with the Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences, China. His research interests include computer vision, space image processing, and deep learning.

TAO CHEN received the Ph.D. degree from the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, China, in 2007. He is currently a Research Fellow and a Supervisor of Ph.D. Candidates at the Chinese Academy of Sciences. His research interests include digital image processing and photoelectric measurement.

JUNCHI LIU received the Ph.D. degree in mechatronic engineering from the University of Chinese Academy of Sciences, in 2016. He is currently an Associate Professor at the Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences. His research interests include space target detection, near-earth asteroid detection, and image processing.

YUAN LIU received the B.S. degree in mechanical engineering from Sichuan University (SCU), Chengdu, China, in 2018. He is currently pursuing the Ph.D. degree in mechanical manufacturing and automation with the Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences, China. His research interests include computer vision, robot visual servo, and deep learning.

QICHANG AN received the B.S. degree from the School of Engineering Science, University of Science and Technology of China, Hefei, Anhui, in 2011, the M.S. degree from the University of Chinese Academy of Sciences, Beijing, in 2014, and the Ph.D. degree from the Changchun Institute of Optics, Fine Mechanics, and Physics, Chinese Academy of Sciences, Changchun, Jilin, in 2018. He is currently a Research Assistant with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His main research interest includes advanced wavefront detection.