

# STAR-24K: A Public Dataset for Space Common Target Detection

**Chaoyan Zhang<sup>1</sup>, Baolong Guo<sup>1\*</sup>, Nannan Liao<sup>1</sup>, Qiuyun Zhong<sup>1</sup>,  
Hengyan Liu<sup>1</sup>, Cheng Li<sup>1</sup> and Jianglei Gong<sup>1,2</sup>**

<sup>1</sup> Institute of Intelligent Control and Image Engineering, Xidian University,  
Taibai Road, Xi'an 710071, Shaanxi, China  
[e-mail: cyzhang\_xidian@qq.com, blguo@xidian.edu.cn, nnliao@stu.xidian.edu.cn  
15122536766@163.com, bingshan45@163.com, licheng812@stu.xidian.edu.cn]

<sup>2</sup> China Academy of Space Technology, Beijing 100094, China  
[e-mail: gongjianglei@stu.xidian.edu.cn]

\*Corresponding author: Baolong Guo

*Received December 6, 2021; revised January 11, 2022; revised February 10, 2022; accepted February 16, 2022;  
published February 28, 2022*

---

## Abstract

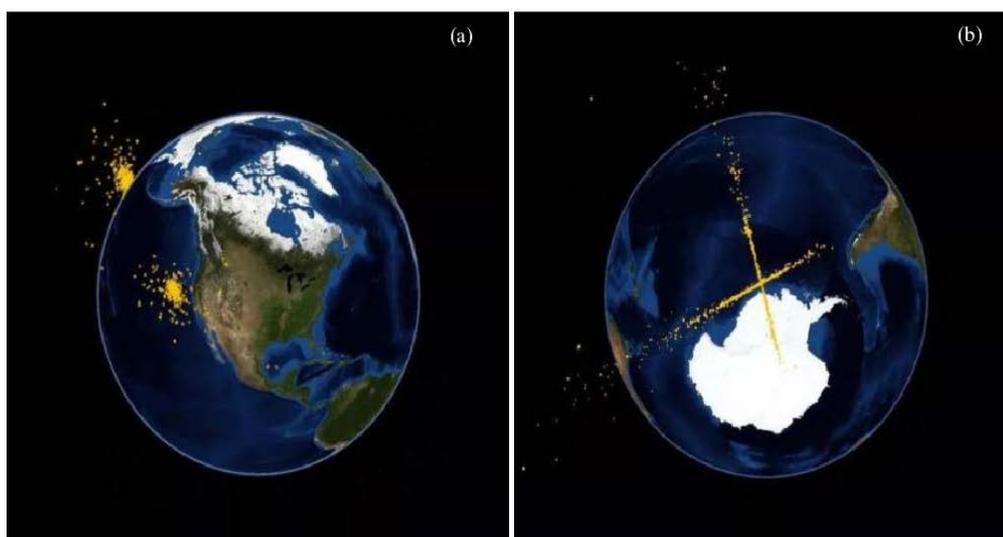
The target detection algorithm based on supervised learning is the current mainstream algorithm for target detection. A high-quality dataset is the prerequisite for the target detection algorithm to obtain good detection performance. The larger the number and quality of the dataset, the stronger the generalization ability of the model, that is, the dataset determines the upper limit of the model learning. The convolutional neural network optimizes the network parameters in a strong supervision method. The error is calculated by comparing the predicted frame with the manually labeled real frame, and then the error is passed into the network for continuous optimization. Strongly supervised learning mainly relies on a large number of images as models for continuous learning, so the number and quality of images directly affect the results of learning. This paper proposes a dataset STAR-24K (meaning a dataset for Space TArget Recognition with more than 24,000 images) for detecting common targets in space. Since there is currently no publicly available dataset for space target detection, we extracted some pictures from a series of channels such as pictures and videos released by the official websites of NASA (National Aeronautics and Space Administration) and ESA (The European Space Agency) and expanded them to 24,451 pictures. We evaluate popular object detection algorithms to build a benchmark. Our STAR-24K dataset is publicly available at <https://github.com/Zzz-zcy/STAR-24K>.

---

**Keywords:** Public dataset, space target detection, deep learning, computer vision.

## 1. Introduction

It has been widely discovered that the recent remarkable success of computer vision technologies, especially those based on deep learning, relies heavily on large-scale and well-annotated datasets. The target detection algorithm based on supervised learning is the current mainstream algorithm for target detection. A high-quality dataset is the prerequisite for the target detection algorithm to obtain good detection performance. The larger the number and quality of the dataset, the stronger the generalization ability of the model, that is, the dataset determines the upper limit of the model learning. For example, CIFAR-10/100 [1] and ImageNet [2] are important activators for deep convolutional neural networks [3]. Pascal VOC [4] and MS COCO [5] are used for common object detection and segmentation, and LFW [6] is used for face recognition. Caltech Pedestrians [7] and MOT benchmark [8] are used for human detection and tracking. Among all these tasks, it is because an excellent dataset promotes the development of various detection technologies. However, due to the sensitivity of satellite and other data, there is no publicly available dataset, evaluating the accuracy of space target detection in optical images is also difficult.



**Fig. 1.** Satellite collision diagram

In recent years, with the rapid development of satellite and rocket technology, more and more spacecraft have entered space. When spacecraft are working, they often encounter satellites and space junk from other countries. If two satellites collided, thousands of pieces of space debris would be generated in space, as shown in Fig. 1, where (a) the distribution of debris 20 minutes after the collision; (b) Debris distribution 50 minutes after the collision. The collision of these generated debris on aircraft will have serious consequences, as shown in Table 1. These objects exist in space and will accumulate more and more, and in the long run, it will definitely affect the development of human aerospace industry and the operation of the earth, so the detection of satellites or debris in space is a very meaningful work. Although deep convolutional networks are very effective for feature extraction and object detection, but the premise of target detection must have a good dataset, and the process of iterative training of

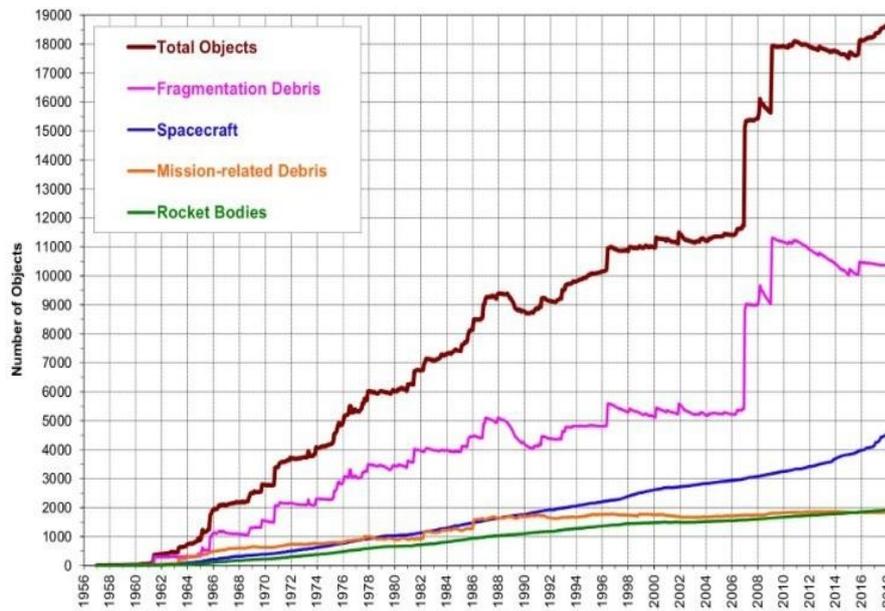
convolutional neural networks is driven by massive amounts of data. At present, for images of space targets, relevant data sets at home and abroad are very scarce. Therefore, this article collected a large number of satellite and space junk images for expansion and annotation, and constructed a data set STAR-24K (meaning Space TArget Recognition with more than 24,000 images), which can be used to detect space targets. It is the labeled dataset that matters in the era of deep learning.

**Table 1.** Space debris colliding experiments on aerospace aluminum

Space debris diameter(mm)	Impact crater depth(mm)	Penetration thickness of aluminum plate(mm)	Perforation diameter(mm)
1	2.17	4.8	4.7
5	11.9	26.1	23.3
10	24.7	44.4	46.5
100	28.1	54.3	465.6

## 2. Related Work

With the increasingly frequent space activities, the number of space targets in the earth's orbit has increased year by year, and it is inevitable that more and more space debris will be brought. The number of cataloged space targets in the Earth's orbit has been increasing year by year, as shown in Fig. 2 [9]. Therefore, the development of space target detection technology has no time to delay.



**Fig. 2.** The increasing situation of the number of space objects on earth orbit with each passing year.

Space target detection is divided into two types: ground-based and space-based. Among them, ground-based detection is mainly composed of ground-based photoelectric detectors, radar detectors, etc. It has the advantages of mature technology, convenient use and maintenance, but due to the limitation of meteorological conditions, geographical location and

other factors [10], there are still many limitations in the detection range and timeliness of the target. Space-based space target detection uses imaging measurement devices installed on the space platform to detect and track targets. It has the advantages of being unaffected by weather and flexible in maneuverability [11], which represents a new research direction in space target detection. Space-based space target visible light imaging process as shown in Fig. 3: the sun, the earth reflected light and moonlight space stray light irradiation, the target after target surface reflection to observation platform, and then through the observation of the CCD sensor system transmission platform and focus, formed in the imaging plane background to black, the goal for the dot light spot observation images. The main task of space target detection is to detect, track, and identify space targets, and determine the shape, size, and movement status of the target. However, many countries' research on space-based space target detection system is relatively late, and there is a big gap between it and foreign countries in target detection and recognition. Therefore, in-depth research on space-based space target detection technology and a large-scale and well-annotated dataset will greatly enhance detection capabilities for space targets.

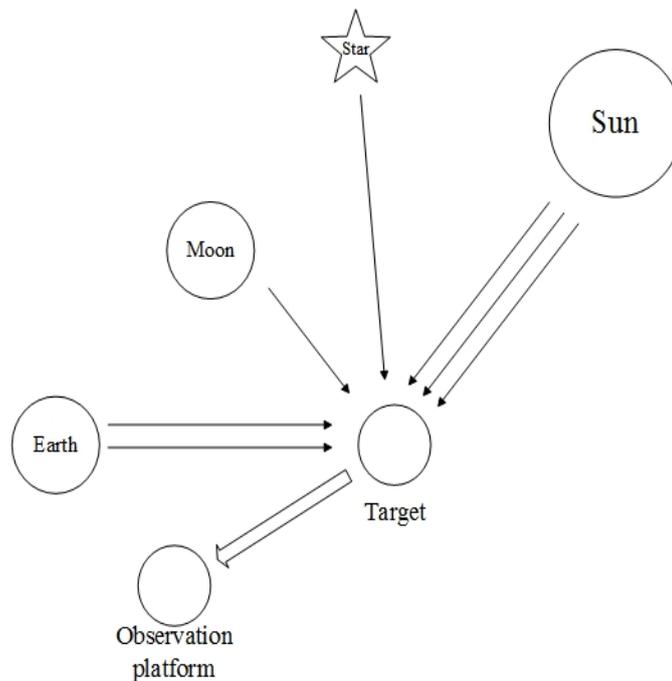


Fig. 3. Schematic diagram of space-based space target imaging.

In many space target detection systems, the use of optical images to detect and track space targets is still a difficult problem. Due to the large distance between the space target and the detector, the target imaging usually only covers a small number of pixels, which means that the target looks like a low-intensity point embedded in strong background clutter [12]. Therefore, although traditional feature-based target detection methods perform well under uniform background conditions, they may not be able to detect space targets under complex backgrounds, such as template matching [13, 14], morphological operations [15, 16, 17], Threshold method [18, 19], high-pass filtering [20], Hessian matrix [21, 22]. In addition, in the field of infrared surveillance [23, 24, 25], a variety of space target detection methods have also

been proposed, but the detection results in space-based visible optical images are not ideal. Reed et al., “[26, 27] proposed a recursive space target detection algorithm based on three-dimensional matched filtering, which theoretically shows the best detection performance for space targets with the same moving speed. However, for targets with unknown speeds, the performance of the algorithm decreases. In 2015, Yao et al., “[28] proposed a space target detection and velocity estimation method based on the maximum projection method. These space target detection algorithms mentioned are not based on deep learning, and many of them use artificially designed target features to achieve single target detection. Then use simulation software such as TASAT, SVST and STK EOIR to simulate the space scene and then carry out the detection task. It can be seen that due to the extreme lack of spatial target datasets, the development of deep learning in the detection of space targets is very slow. Therefore, proposing a dataset on space targets is of great significance to the development of this field. After our investigation, there are three verified data sets about satellite targets, BUAA-SID [29], SPEED [30] and SPEED+ [31]. The BUAA-SID dataset was established in 2010, and the satellite images in the dataset were all simulated by computer rather than real optical images, which lacked a certain scientific nature. The amount of data required for deep learning is also relatively small. SPEED and SPEED+ data sets are both used to solve the satellite pose estimation problem, not to perform satellite detection tasks. However, in the field of space target detection, there is still a lack of data set, so the data set we established is relatively leading in both the quantity and quality of images. While challenging and costly, it is very necessary to establish a large-scale benchmark to facilitate research in this field.

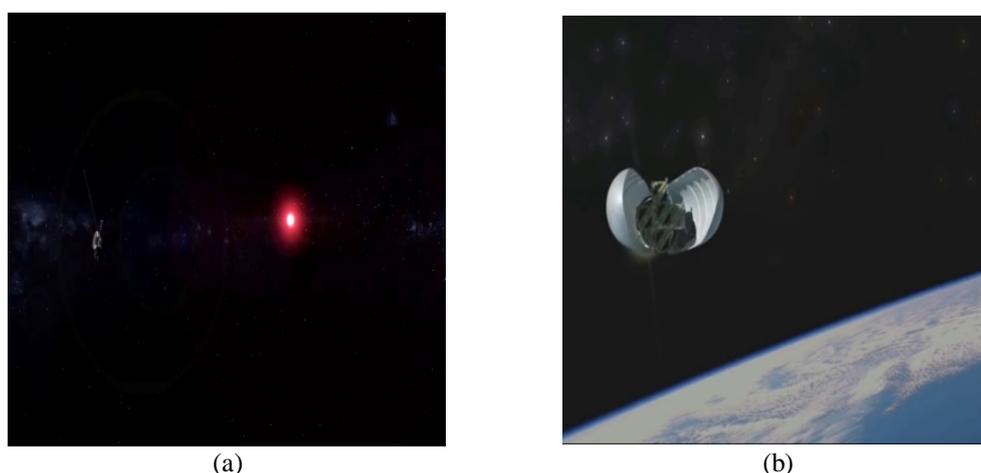
### 3. Establishment of STAR-24K

#### 3.1 Images Collection

Obtaining images is the first step in dataset production. Since there are fewer tasks related to space target detection and have a certain degree of confidentiality. However, target detection requires a large number of samples to train the detector, and it is very difficult to obtain the required number of datasets for target detection tasks, so it is necessary to manually make a dataset for space target detection. Most of the pictures about space junk, satellites, and planets collected directly from search engines are too illusory and not real enough, and some are of low quality so that there are few pictures that really meet the requirements of dataset production. Therefore, the pictures of the data set produced in this article are mainly derived from three aspects: (1) Select from the pictures published by NASA (National Aeronautics and Space Administration) and ESA (The European Space Agency) on the official websites, and take a frame-by-frame screenshot from the video they released about space exploration. (2) Single-frame interception from short space exploration science videos published by a series of internet video sites such as YouTube. (3) Watch "Space Junk", "Horizon" and other documentaries, from which intercept images related to space satellites, junk, and planets.

In the process of video interception, this article specially selected different angles, different lights, different positions, different scenes and different types of pictures to ensure the diversity of the targets in the pictures and helps prevent under-fitting and over-fitting during the training process. The pictures selected in this article mainly include three categories: satellites, planets, and space junk (as shown in Fig. 4). The three types of targets selected are also the most common objects in space. However, because there are not many resources on space target detection on the Internet, the total number of collected images is 600, which

cannot reach the number required for network learning. Therefore, this article first mixes the collected pictures as the original picture, and then uses data augmentation methods to amplify, the amplified pictures reach 24,451. For general network training, the number of this dataset meets the requirements.



**Fig. 4.** Some common space targets. (a) On the left is an artificial satellite, on the right is a planet, (b) represents space junk, which is generally unregulated.

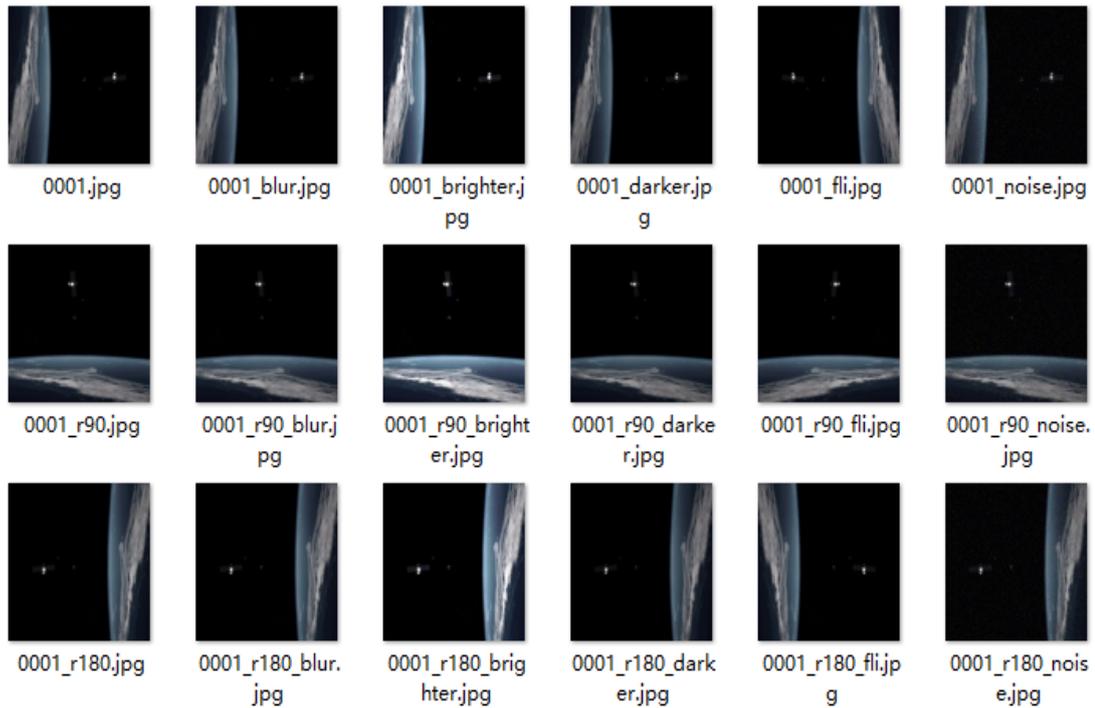
### 3.2 Data Augmentation

Data augmentation technology is the most effective method to solve the problem of small number of original data sets. Data augmentation is a general term for methods to expand data. It can increase the number of samples in the data set, effectively alleviate the problem of model overfitting, and can also bring better generalization capabilities to the model. There are two ways of data augmentation. One is offline augmentation. Offline augmentation refers to performing all the augmentations in advance, and then sending them to the network for learning. This method is suitable for the case where the data set is relatively small, and the number of pictures is doubled several times through methods such as rotation, translation, and flipping. Another data augmentation method is online augmentation. Online augmentation is to send the original data directly to the network for small batch conversion. This method is suitable for large data sets.

This article chooses offline data augmentation. Offline data augmentation can be divided into pixel data augmentation and image data augmentation. Pixel data augmentation includes changing brightness and adding noise. Image data augmentation mainly includes cropping, translation, mirroring, rotation, and occlusion. The methods used in the data augmentation in this article include darkening, brightening, blurring, adding Gaussian noise and salt and pepper noise, mirroring, turning the original image by 90 degrees and turning it by 180 degrees. Then flip 90 degrees and 180 degrees on the darkened, brightened, blurred, and increased noise pictures, and then change the brightness and increase the noise on the flipped pictures (as shown in Fig. 5), and finally make the data set on the original basis through these operations increased by 17 times. Google Brain [32] Team proved that deep learning can achieve strong performance on large amounts of data, and data augmentation is required to increase the training data of deep learning. Scale change, translation and rotation are often used to enhance data amount. In addition, it is also effective to randomly add noise and occlusion to the image. At the same time, some dataset papers also use data augmentation when establishing data sets,

such as Yanghua Di et al., “s ship dataset FGSCR-42 dataset [33].

Then use the enhanced data for image stitching, and combine the four images into one, which greatly enriches the background and target of the detected object, and because of the distance, many targets in space are small targets. Therefore, splicing in this way also provides help for the detection of small targets and increases the robustness of the model. Sending this new picture to the network for learning is equivalent to changing from only learning one picture at a time to four pictures at a time, which greatly improves the learning ability of the model.



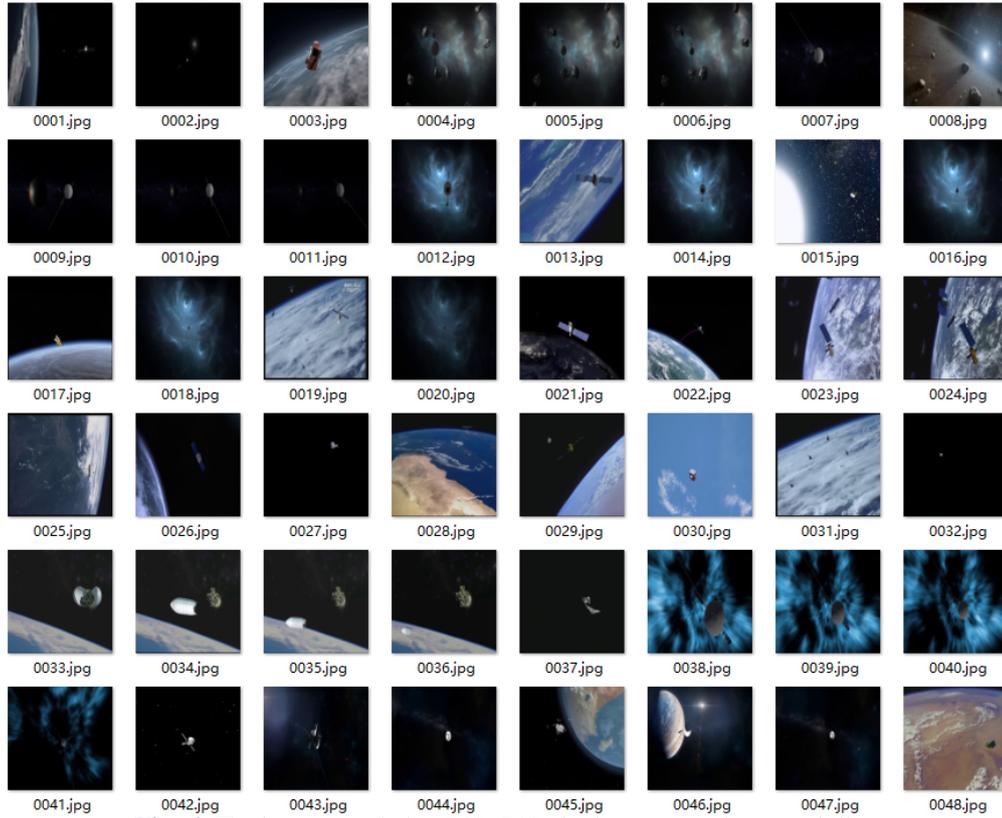
**Fig. 5.** Data augmentation by darkening, brightening, blurring, adding Gaussian noise, etc.

**Table 2.** The number of instances of each class

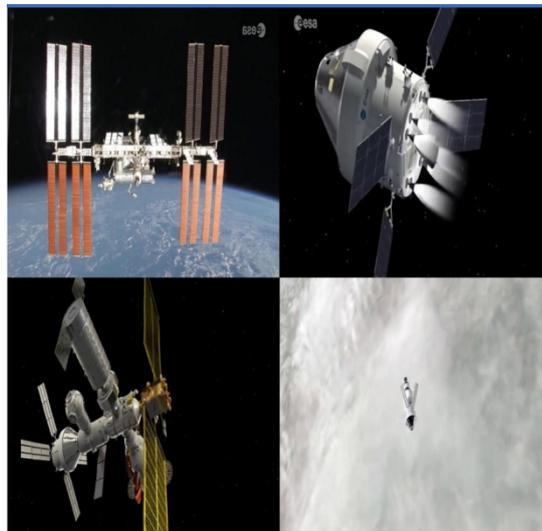
Name of class labels	Number of instances in training set
Satellite	39892
Stars	4088
Space Junk	157
Total	44137

Before data augmentation, the original pictures are named in order, and then the pictures are uniformly scaled to  $960 \times 960$  pixels, as shown in Fig. 6. Then the pictures are augmented, and the augmented pictures are 17 times larger than the original. After the pictures are augmented, the pictures are seamlessly spliced, and four pictures are randomly selected to form one picture. The picture pixels after splicing are  $1920 \times 1920$  pixels, as shown in Fig. 7. A total of 600 pictures were collected, and the pictures were divided into 480 pictures in the training set and 120 pictures in the test set. Then use these 600 pictures as the original pictures for data augmentation. Separating the original pictures first and then performing data augmentation can ensure that there is no same picture in the training set and the test set. This

can effectively prevent data leakage in the process of training and testing from affecting the performance of the trained model. The data set after data augmentation is 24451, among which the training set is 20451 and the test set is 4000. The number of instances of each class are shown in [Table 2](#). In this data set, we mainly focus on detecting satellites, so satellites account for a large proportion of instances.



**Fig. 6.** The images scaled to  $960 \times 960$  pixels are named sequentially.



**Fig. 7.** Randomly select four pictures for stitching.

### 3.3 Image Annotation

Since the data set is the self-selected images and needs to be manually labeled, this article uses LabelImg software for manual labeling. To obtain high-quality space target annotations for each image, we recruit 10 well-trained annotators and ask them to annotate all the space targets they can distinguish. Three rounds of cross-checking and correction are then carried out to improve the annotation quality. Finally, it takes about one month to complete the whole annotation process, where 20451 images are labeled. See Fig. 8 for example images with annotation.

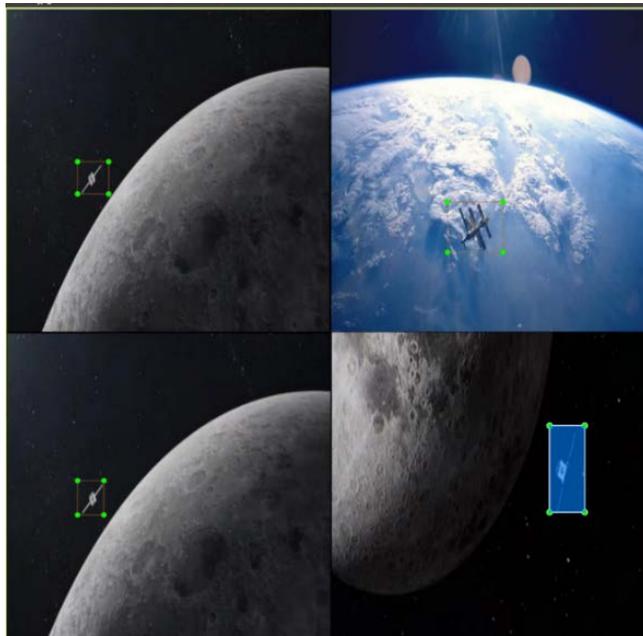


Fig. 8. Example images with annotation.

Considering the ease of sharing and saving to the computer, the annotations are saved in line with the COCO format to facilitate further research in space target detection by reusing common training and evaluation tools developed in the object detection community. The STAR-24K is split into three disjoint subsets, i.e., train, validation, and test sets, at the ratio of 8:1:1 per space target species.

## 4. Evaluations

Common target detection networks can be used to detect space targets, but the detection results need to be improved. Considering that the space target detection task requires high real-time performance, there are also certain requirements for detection accuracy. For the comprehensiveness of the results, we eventually selected CenterNet [34], YOLOv3 [35], YOLOv4-tiny [36], YOLOv4 [37] and YOLOv5 [38], as our testing algorithms for building a benchmark on STAR-24K. These baseline models cover the popular and the state-of-the-art deep learning target detection networks [39] [40] [41], thus can offer valuable benchmark results for dataset evaluation. Meanwhile, to demonstrate the effectiveness of data

augmentation, we add a set of comparative experiments. After data augmentation, the experiment is to split the data set according to 8:2. As a result, the 600 initial images were divided in a ratio of 8:2. Among them, 480 are used for training, and 120 are used for testing. Compared with the training on YOLOv5 algorithm, the validity of data augmentation is proved (as shown in **Table 3**).

**Table 3.** Comparison of initial data and data augmentation

Type of Data	mAP <sub>.50</sub>
Initial Data	16.8%
Data Augmentation	82.9%

For the task of space target detection, the performance is evaluated as the recall and mAP.50 as metrics. Average precision (AP) refers to the area enclosed under the PR curve, and mean average accuracy (mAP) refers to the average accuracy rate under different categories, which is to calculate AP values for multiple verification sets; Recall refers to the proportion of actual positive samples in all prediction samples relative to samples. Their calculation formula is as follows. mAP.50 is the mean average precision at IoU = 0.50.

$$AP = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Where TP indicates that the actual value is a positive sample, and the predicted value is also a positive sample number; TN indicates that the actual value is a negative sample, and the predicted value is also a negative sample number; FP refers to the number of samples with negative actual value and positive predicted value; FN indicates that the actual value is a positive sample and the predicted value is a negative sample.

$$mAP = \frac{\sum_{i=1}^n AP(i)}{n} \quad (3)$$

Where  $i$  is the current category,  $n$  is the number of categories, and  $AP(i)$  is the accuracy of the current category.

**Table 4.** The detection result of baseline model

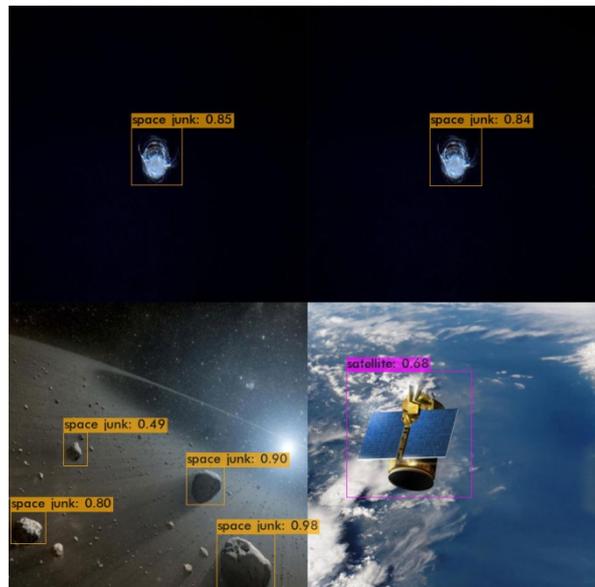
Model	mAP <sub>.50</sub>	Recall
CenterNet	30.9%	35.3%
YOLOv3	74.3%	81.1%
YOLOv4-tiny	73.97%	81.0%
YOLOv4	78.5%	87.0%
YOLOv5	82.9%	96.1%

This experiment is conducted in the environment of Ubuntu 16.04 operating system, 16GB memory and Titan x graphics card, then the network weight is obtained through training, and the weight is used to test the pictures of the test set. Benchmark results of the selected baseline

models are shown in **Table 4**. It can be seen that the recall is higher than the other index. Since the number of artificial satellites in the data set is more than that of other targets, the recall will be higher, but the mAP.50 is in the normal range. **Fig. 9**, **Fig. 10**, **Fig. 11** and **Fig. 12** are the test results of algorithms CenterNet, YOLOv4-tiny, YOLOv4 and YOLOv5 respectively from an image selected from dataset STAR-24K. The test results also shows that YOLOv5 performs better.



**Fig. 9.** CenterNet test result.



**Fig. 10.** YOLOv4-tiny test result.



development of space target detection algorithms in optical images. In the future work, we will further expand the amount of categories and the size of the dataset. If possible, we will invite more experts to make more precise annotations of the dataset.

## Acknowledgement

Thanks to the editor and anonymous reviewers for their suggestions and comments that helped improve the final version of the paper.

## References

- [1] Antonio Torralba, Robert Fergus, and William T. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(11), 1958–1970, 2008. [Article \(CrossRef Link\)](#)
- [2] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, 115(3), 211–252, 2015. [Article \(CrossRef Link\)](#)
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, 60(6), 84–90, 2017. [Article \(CrossRef Link\)](#)
- [4] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision*, 88(2), 303–338, 2010. [Article \(CrossRef Link\)](#)
- [5] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick, "Microsoft coco: Common objects in context," in *Proc. of European conference on computer vision*, pp. 740–755, 2014. [Article \(CrossRef Link\)](#)
- [6] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," 2008. [Online]. Available: [https://www.researchgate.net/publication/29622837\\_Labeled\\_Faces\\_in\\_the\\_Wild\\_A\\_Database\\_for\\_Studying\\_Face\\_Recognition\\_in\\_Unconstrained\\_Environments](https://www.researchgate.net/publication/29622837_Labeled_Faces_in_the_Wild_A_Database_for_Studying_Face_Recognition_in_Unconstrained_Environments)
- [7] Piotr Dollár, Christian Wojek, Bernt Schiele, and Pietro Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE transactions on pattern analysis and machine intelligence*, 34(4), 743–761, 2012. [Article \(CrossRef Link\)](#)
- [8] Anton Milan, Laura Leal-Taixé, Ian Reid, Stefan Roth, and Konrad Schindler, "Mot16: A benchmark for multi-object tracking," *arXiv*, 2016. [Article \(CrossRef Link\)](#)
- [9] NASA. Orbital Debris Quarterly News [EB/OL], 2017-2. [Online]. Available: <https://directory.eoportal.org/web/eoportal/satellite-missions/content/-/article/orbital-debris>
- [10] Wirnsberger H, O Baur, G Kirchner, "Space Debris Orbit Prediction Errors Using Bi-Static Laser Observations," *Adv. Space Res.*, 55(11), 2607–2615, 2015. [Article \(CrossRef Link\)](#)
- [11] Nunez J, A Nunez, F J Montojo, et al., "Improving Space Debris Detection in GEO Ring Using Image Deconvolution," *Adv. Space Res.*, 56(2), 218–228, 2015. [Article \(CrossRef Link\)](#)
- [12] Sease B, B Flewelling, J Black, "Automatic Streak Endpoint Localization from The Cornerness Metric," *Acta Astronautica*, 134, 345–354, 2017. [Article \(CrossRef Link\)](#)
- [13] Liu R, Li X, Han L, et al., "Track infrared point targets based on projection coefficient templates and non-linear correlation combined with Kalman prediction," *Infrared Phys. Technol.*, 57, 68–75, 2013. [Article \(CrossRef Link\)](#)
- [14] Bal A, Alam M S, "Automatic target tracking in FLIR image sequences using intensity variation function and template modeling," *IEEE Trans. Instrum. Meas.*, 54(5), 1846–1852, 2005. [Article \(CrossRef Link\)](#)

- [15] Bai X, “Morphological operator for infrared dim small target enhancement using dilation and erosion through structuring element construction,” *Optik*, 124(23), 6163-6166, 2013. [Article \(CrossRef Link\)](#)
- [16] Wei M, Xing F, You Z, “A real-time detection and positioning method for small and weak targets using a 1D morphology-based approach in 2D images,” *Light: Science & Applications*, 7, 2018. [Article \(CrossRef Link\)](#)
- [17] Bai X, Zhou F, Xie Y, et al., “Enhanced detect ability of point target using adaptive morphological clutter elimination by importing the properties of the target region,” *Signal Process*, 89(10), 1973-1989, 2009. [Article \(CrossRef Link\)](#)
- [18] Chang Y K, Kang S J, Lee B H, “High-accuracy image centroiding algorithm for CMOS-based digital sun sensors,” in *Proc. of SENSORS, 2007 IEEE*, 329-336, 2007. [Article \(CrossRef Link\)](#)
- [19] Xu W, Li Q, Feng H J, et al., “A novel star image thresholding method for effective segmentation and centroid statistics,” *Optik*, 124(20), 4673-4677, 2013. [Article \(CrossRef Link\)](#)
- [20] Yang L, Yang J, Yang K, “Adaptive detection for infrared small target under sea-sky complex background,” *Electron. Lett.*, 40(17), 1083-1085, 2004. [Article \(CrossRef Link\)](#)
- [21] Chae S W, Kwon K Y, Lee T S, “An optimal design system for spot welding locations,” *Finite Elements Anal. Des.*, 38(3), 277-294, 2002. [Article \(CrossRef Link\)](#)
- [22] Jiang J, Deng J, Zhang G, et al., “Real-time implementation of light stripe center extraction,” *Opt. Tech.*, 34, 170-173, 2008. [Article \(CrossRef Link\)](#)
- [23] Gao J, Lin Z, An W, “Infrared Small Target Detection Using a Temporal Variance and Spatial Patch Contrast Filter,” *IEEE Access*, 7, 32217-32226, 2019. [Article \(CrossRef Link\)](#)
- [24] Zhang T, Cui C, Fu W, et al., “Improved Small Moving Target Detection Method in Infrared Sequences under A Rotational Background,” *Applied optics*, 57(31), 9279-9286, 2018. [Article \(CrossRef Link\)](#)
- [25] Chen Y, “Infrared Small Target Detection Through Multiple Feature Analysis Based on Visual Saliency,” *IEEE Access*, 7, 38996-39004, 2019. [Article \(CrossRef Link\)](#)
- [26] Reed I S, Gagliardi R M, Stotts L B, “A Recursive Moving-Target-Indication Algorithm for Optical Image Sequences,” *IEEE Trans. Aerosp. Electron. Syst.*, 26(3), 434-440, 1990. [Article \(CrossRef Link\)](#)
- [27] Reed L, Gagliardi R, Stotts L, “Optical moving target detection with 3-D matched filtering,” *IEEE Transactions on Aerospace and Electronic Systems*, 24(4), 327-336, 1988. [Article \(CrossRef Link\)](#)
- [28] Yao D, Wen D, Xue J, et al., “Maximum Projection and Velocity Estimation Algorithm for Small Moving Target Detection in Space Surveillance,” *Image Processing and Analysis*, 96752H, 2015. [Article \(CrossRef Link\)](#)
- [29] Zhang H, et al., “BUAA-SID1.0 Space object image database,” *Space Return and Remote Sensing*, 31(04), 65-71, 2010. [Article \(CrossRef Link\)](#)
- [30] Kisantal M, Sharma S, Park T H, et al., “Satellite Pose Estimation Challenge: Dataset, Competition Design and Results,” *arXiv*, 2019. [Article \(CrossRef Link\)](#)
- [31] Tae H, Marcus M, Gurvan L, et al., “SPEED+: Next Generation Dataset for Spacecraft Pose Estimation across Domain Gap,” *arXiv*, 2021. [Article \(CrossRef Link\)](#)
- [32] Zoph B, Cubuk E D, Ghiasi G, et al., “Learning Data Augmentation Strategies for Object Detection,” *arXiv*, 2019. [Article \(CrossRef Link\)](#)
- [33] Di Y, Jiang Z, Zhang H, “A Public Dataset for Fine-Grained Ship Classification in Optical Remote Sensing Images,” in *Proc. of Image and Signal Processing for Remote Sensing XXV*, 2019. [Article \(CrossRef Link\)](#)
- [34] Zhou, X., Wang, D., and Krahenbuhl, P, “Objects as points,” *arXiv: 1904.07850*, 2019. [Article \(CrossRef Link\)](#)
- [35] Redmon J, Farhadi A, “YOLOv3: An Incremental Improvement,” *arXiv e-prints*, 2018. [Article \(CrossRef Link\)](#)
- [36] Alexey AB. [Online]. Available: <https://github.com/AlexeyAB/darknet>, 2020.
- [37] Alexey Bochkovskiy, Chien-Yao Wang, and HongYuan Mark Liao, “Yolov4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv: 2004.10934*, 2020. [Article \(CrossRef Link\)](#)

- [38] Glenn Jocher et al., "yolov5". [Online]. Available: <https://github.com/ultralytics/yolov5>, 2021.
- [39] Zaidi, S. S. A., Samar Ansari, M., Aslam, A., Kanwal, N., Asghar, M., and Lee, B., "A Survey of Modern Deep Learning based Object Detection Models," *arXiv e-prints*, 2021.  
[Article \(CrossRef Link\)](#)
- [40] L. Jiao et al., "A Survey of Deep Learning-Based Object Detection," *IEEE Access*, vol. 7, pp. 128837-128868, 2019, [Article \(CrossRef Link\)](#).
- [41] Chen, S. -H., Wang, C. -W., Tai, I. -H., Weng, K. -P., Chen, Y. -H., & Hsieh, K. -S., "Modified YOLOv4-DenseNet Algorithm for Detection of Ventricular Septal Defects in Ultrasound Images," *International Journal Of Interactive Multimedia And Artificial Intelligence*, 6(Special Issue on Current Trends in Intelligent Multimedia Processing Systems), 101-108, 2021.  
[Article \(CrossRef Link\)](#)



**CHAOYAN ZHANG** received the B.S. degree in aerospace science and technology major from the School of Aerospace Science and Technology, Xidian University, Xi'an, China, in 2018. He is currently pursuing doctor's degree with the School of Aerospace Science and Technology at Xidian University. His main research interests include deep learning and neural networks.



**BAOLONG GUO** received the B.S., M.S., and Ph.D. degrees from Xidian University, in 1984, 1988, and 1995, respectively, all in communication and electronic systems. From 1998 to 1999, he was a Visiting Scholar with Doshisha University. He is currently a Full Professor with the School of Aerospace Science and Technology, Xidian University, Xi'an, China. His current research interests include pattern recognition, intelligent information processing, image processing, and video communication. He has published over 100 academic papers and organized several international conferences as the Co-Chair.



**QIUYUN ZHONG** received the bachelor's degree in automation from Civil Aviation University of China in 2019. Now she is studying for her master's degree in control science and engineering from Xi'an University. During her master's degree, her main research direction is object detection based on infrared image.



**HENGYAN LIU** received the B.S. degree in automation from QiLu University of Technology, Shandong, China, in 2019. He is currently pursuing the M.S. degree in Control engineering from Xidian University, Shaanxi, China. His research interest includes the development of image processing, deep learning, and the text detection and recognition process.



**NANNAN LIAO** received the B.S. degree in Communication engineering major from Xi'an Shiyou University (Xi'an, China), in 2018. She is currently pursuing doctor's degree with the School of Aerospace Science and Technology at Xidian University. Her main research interests include image processing and information processing.



**CHENG LI** received the B.S. degree from the School of Electro-mechanical Engineering, Xidian University, Xi'an, China, in 2013, and the M.S. degree from the School of Aerospace Science and Technology, Xidian University, in 2017, where he is currently pursuing the Ph.D. degree with the Intelligent Control and Image Engineering Institute. He focuses on superpixel-based video image representation and its applications.



**JIANGLEI GONG** was born in 1986. He received the B.E. and M.S. degrees from Beihang University in 2008 and 2011 respectively. He is currently working in China academy of space technology and studying at Intelligent Control and Image Engineering Institute of Xidian University for his doctor's degree. His main research interests is the design of communication satellite information system.