



NEA Detection Method with Neural Network in Sidereal Tracking

Yijun Tang^{1,2}, Yunxiao Jiang^{1,2}, Zhen Zhang^{1,2}, Chenchen Ying^{1,2}, Songqi Zhang^{1,2}, Liangcheng Liao^{1,2}, Junjie Ma^{1,2}, Bo Yan^{1,2}, Chunhai Bai³, Guojie Feng³, Xiaoming Zhang⁴, and Xiaojun Jiang⁴

¹ School of Physics, Zhejiang University of Technology, Hangzhou 310023, China; tyj1970@zjut.edu.cn

² Collaborative Innovation Center for Bio-Med Physics Information Technology of ZJUT, Zhejiang University of Technology, Hangzhou 310023, China

³ Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011, China

⁴ CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

Received 2025 March 10; revised 2025 May 9; accepted 2025 June 6; published 2025 July 10

Abstract

Near-Earth Asteroids posed a threat to human civilization, making their monitoring crucial. As the demand for asteroid detection technology increased, precise detection of these celestial bodies became an urgent task to understand their characteristics and assess potential impact risks. To improve asteroid detection accuracy and efficiency, we proposed an advanced image processing method and a deep learning network for automatic asteroid detection. Specifically, we aligned star clusters and overlaid images to exploit asteroid motion rates, transforming them into object-like trajectories and improving the signal-to-noise ratio. This approach created the Asteroid Trajectory Image Data set under various conditions. We modified CenterNet2 network to develop AstroCenterNet by integrating Multi-channel Histogram Truncation for feature enhancement, using the SimAM attention mechanism to expand contextual information and suppress noise, and refining Feature Pyramid Network to improve low-level feature detection. Our results demonstrated a detection accuracy of 98.4%, a recall of 97.6%, a mean Average Precision of 94.01%, a false alarm rate of 1.6%, and a processing speed of approximately 17.86 frames per second, indicating that our method achieves high precision and efficiency.

Key words: methods: data analysis – techniques: image processing – minor planets, asteroids: general – planets and satellites: detection

1. Introduction

Asteroids are small celestial bodies in the Solar System that orbit the Sun, similar to planets but significantly smaller in volume and mass. Most of the asteroids are concentrated in the asteroid belt located between the orbits of Mars and Jupiter. These objects are generally considered remnants of planetesimals from the early stages of Solar System formation and represent the most numerous class of objects within the Solar System. Studying asteroids can help us understand information about the origin of life and the potential threats they pose to the Earth (Azadmanesh et al. 2023). Due to the proximity of asteroids to Earth and the potential for impact, they have long been a significant concern in the field of planetary defense. An example of this danger occurred on 2023 February 12. The Piskéstető Observatory in Hungary discovered asteroid 2023 CX1. In response to growing concerns about asteroid impacts, various countries have developed strategies and technologies to detect, track, and mitigate potential threats. The United States, through NASA's Planetary Defense Coordination Office, has been conducting several space missions such as the mission to test the Double Asteroid Redirection Test, which aims to test the effectiveness of using kinetic impactors to alter the trajectory of potentially hazardous asteroids. On the European side, recognizing the need for asteroid monitoring,

the European Space Agency initiated the Ramses mission, aimed at exploring the asteroid Apophis. This mission focuses on detailed studies of Apophis to assess its potential risk to Earth and to advance scientific understanding of asteroids. Ramses will utilize various scientific instruments for observation and data collection, providing crucial information for developing future defense strategies. China has also been actively involved in the observation and tracking of asteroids. One notable project is the wide-field survey program carried out by the Xinjiang Astronomical Observatory. This program uses wide-field optical telescopes to observe and track Near-Earth Objects (NEOs) (Bai et al. 2020), including asteroids, providing valuable data for planetary defense efforts.

The initial step in asteroid studies often involves the detection and search for asteroids. In wide-field telescope images, multiple moving objects may appear within a single field of view, exhibiting varying speeds and directions (Virtanen et al. 2016). Processing astronomical images is more complex than conventional image processing because many astronomical objects lack clear boundaries and are intertwined with noise components. Furthermore, when displaying deep wide-field-of-view images from multiple sources, the sizes and intensities of different objects can vary significantly (Masias et al. 2012). One of the key challenges in

processing survey images is managing the large volume of data, as these images typically encompass vast areas of the sky. The large number of objects captured in a single image, combined with inherent noise from natural sources (such as cosmic radiation) and instrumental factors (such as sensor imperfections), complicates the task of distinguishing relevant asteroids from other celestial bodies and artifacts. Additionally, asteroids often manifest as faint streaks or points, making them difficult to differentiate from background noise, particularly in low-resolution or long-exposure images. Furthermore, for moving objects, motion blur caused by their relative velocity adds complexity to accurately determining their location and trajectory. Consequently, object detection algorithms play a critical role in identifying asteroid objects in astronomical images. For both ground-based and space-based telescope observations, robust detection algorithms are essential for identifying potential asteroids.

Most traditional asteroid detection algorithms are based on image processing and filtering methods. For example, Bertin & Arnouts (1996) developed SExtractor, utilizing thresholding and connected domain analysis, achieving over 95% reliability in star and galaxy separation. Further, Hobson & McLachlan (2003) proposed a Bayesian framework for detecting and characterizing discrete objects in noisy astronomical data, the approach demonstrated its reliability in identifying objects embedded in diffuse backgrounds; Pan et al. (2014) proposed a detection method for small and dim space targets in time series images using adaptive background noise suppression, effectively detecting targets with a signal-to-noise ratio (SNR) as low as 3; while Sun et al. (2015), Sun & Zhao (2013) introduced a morphological approach and median filtering algorithm for detecting space debris, effectively enhancing the detection of faint objects and validating its reliability through tests on CCD images. Pradhan et al. (2019) employed a zenith-pointing telescope in time-delay integration mode to detect and estimate the sizes of space debris, demonstrating its effectiveness in identifying faint debris as small as 50 cm in low Earth orbits. Liu et al. (2020a) proposed a geometric dual multi-object detection method, the method achieves high accuracy in cluttered environments and validates its efficiency on real-world data sets. Liu et al. (2020b) proposed the Spatiotemporal Pipeline Multistage Hypothesis Testing method for detecting space targets in complicated wide-field surveillance situations. By combining spatiotemporal filtering and hypothesis testing, the method effectively eliminates background interference and achieves high detection accuracy. Ohsawa (2022) developed tracee, an algorithm using k-nearest neighbor graphs and line segment grouping to detect moving objects, demonstrating robust performance in handling noisy and cluttered data. Lastly, Jiang et al. (2022) used a three-frame wavelet transform and bootstrap filtering for stellar interference elimination, validated through real-world astronomical data. However, they all share

some common limitations, such as sensitivity to parameter selection, difficulty in detecting small and weak targets under low SNR conditions, and difficulty in coping with fluctuations in background noise, which make them less adaptable in practical applications.

In recent years, due to the rise of deep learning methods, many scholars have used neural networks for space object detection tasks, including asteroids. Petit et al. (2004) integrated into a complete moving object detection package, which ensures high detection accuracy while accelerating efficiency. Xi et al. (2020) utilized a LeNet5 Convolutional Neural Network (CNN) trained on numerous preprocessed simulated streaks with varying SNR and motion parameters to effectively detect low-SNR space debris streaks. He et al. (2021) built a deep learning source detection network based on the YOLOv4 object detection framework to detect sources, and designed a new deep learning classification network called astronomy photometric source classification network to classify sources. Dai et al. (2022) used the PP-YOLOv2 (Huang et al. 2021) algorithm to detect space objects in single frame images, and designed a postprocessing algorithm called CFS for further candidate filtering and supplementation. Cowan et al. (2023) designed and developed several custom CNN models, combined with YOLOv4 (Bochkovskiy et al. 2020), to identify images with asteroid trajectories and locate asteroids. Pontinen et al. (2023) combined a convolutional neural network, a recurrent neural network, and a gradient-boosting tree to construct a deep learning pipeline to detect asteroid fringes in space telescope images. Pan et al. (2024) proposes an end-to-end deep learning supernova detection network, the Real-Time SuperNova DETection TRansformer, which solves the problem of traditional supernova detection pipelines relying on complex and relatively inefficient image differencing techniques by integrating image differencing, source detection, and Real-bogus classification. By using two independent detection algorithms and combining the results, Du et al. (2024) proposed an asteroid tracking algorithm called DSAT, which first converts single frame object detection into a segmentation problem to more robustly extract dim potential moving objects, and then proposed a moving object tracking algorithm based on a multi-frame motion prior to identify real moving objects. Day et al. (2024) proposed an orbital resident space object detection method that utilizes CNNs and edge computing devices to identify direct orbital objects, greatly reducing the requirement for ground processing bandwidth. However, this method has only been validated on simulated star field images, and its effectiveness in detecting real objects requires further research. Wang et al. (2024) extracted optical features of objects in the image to distinguish between the objects, used neural network training to obtain classification results, and finally used morphometric statistics to detect objects with different motion patterns.

Although an increasing number of deep learning-based detection methods have been applied to asteroid detection,

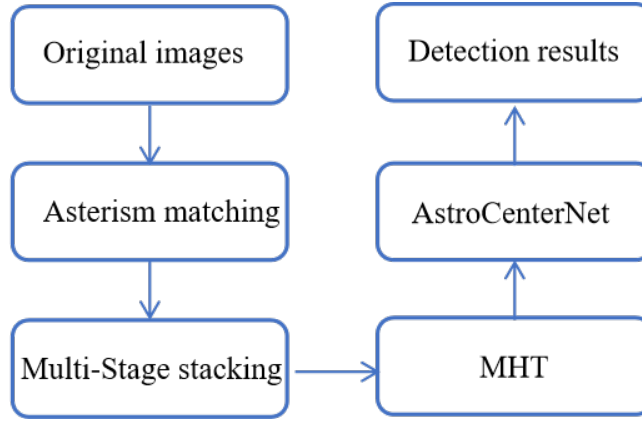


Figure 1. Image processing flow in this article.

effectively training neural networks requires a substantial amount of annotated astronomical image data. In real-world observations, the precision and accuracy of detecting moving celestial objects are influenced by various factors, including complex sky backgrounds, dense star fields, uncertain telescope pointing, non-uniform imaging systems, and optical system distortions (Sun et al. 2019). Moreover, astronomical images captured by telescopes frequently suffer from significant interference due to sky background noise, faint asteroid signals, and discontinuous motion trajectories. These issues lead to a high incidence of false alarms and missed detections in existing detection methods. This paper presents our Asteroid Trajectory Image Data set (ATID), which is constructed from astronomical image data captured by a wide-field telescope operating in the survey mode. In addition to the ATID, we introduce targeted improvements to the object detection algorithm tailored for asteroid identification. To provide a comprehensive overview of our approach, the workflow of the proposed method is illustrated in Figure 1. The subsequent sections offer a detailed description of each component of our methodology:

(1) Image Data Preprocessing: The collected FITS image data are first cropped to focus on regions of interest. Subsequently, Asterism matching (Beroiz et al. 2020) is applied to align and stabilize the background stars, facilitating the identification of asteroids. Asterism matching technology can correct the offsets and distortions caused by camera movement, Earth’s rotation or minor misalignments between images. The aligned images are subsequently stacked multiple times to transform the asteroids into well-defined trajectory objects, thereby reducing interference from static elements like background stars and constructing the data set necessary for neural network training.

(2) Network Architecture and Improvements: Building upon the CenterNet2 framework, we introduce a novel architecture, AstroCenterNet, and integrate a Multi-channel Histogram Truncation (MHT) preprocessing module to enhance image contrast. In addition, the Simple Parameter-Free Attention Module (SimAM)

attention mechanism is integrated to capture contextual information and mitigate the impact of noise. Furthermore, the Feature Pyramid Network (FPN) module is refined by incorporating low-resolution feature fusion and low-resolution feature map detection branches to reduce the false alarm rate (FAR).

(3) Experiments and Evaluation: The improved neural network is validated through the analysis of precision-recall (P-R) curves and loss function graphs. Comparative experiments and ablation studies are conducted using various object detection algorithms to evaluate the accuracy and performance of the network model.

2. NOWT: A Wide-field Observatory for NEOs Survey and Exploration

The Nanshan One-meter Wide-field Telescope (NOWT; Bai et al. 2020) is located at Nanshan Station ($87^{\circ}10'30''$ East, $43^{\circ}28'25''$ North; at an altitude of about 2088 m) of Xinjiang Astronomical Observatory (XAO). The relevant parameters of the telescope and its detectors are shown in Table 1.

3. Data Set and Preprocessing

3.1. Data Collection and Image Cropping

The NOWT was employed as the observational instrument for data acquisition and data set construction. The raw image files are stored in 16 bit depth FITS format, with a resolution of 4096×4096 pixels, as exemplified in panel (a) of Figure 2. The NOWT captures a wide field of celestial objects, which is often subject to interference from moonlight and cloud cover, leading to an uneven distribution of the background greyscale. Applying global greyscale normalization across the entire image would exacerbate the issue by further dimming regions with inherently lower greyscale values. To address this, the original image was segmented into 16 smaller sub-images, each with a resolution of 1048×1048 pixels, as illustrated in panel (b) of Figure 2.

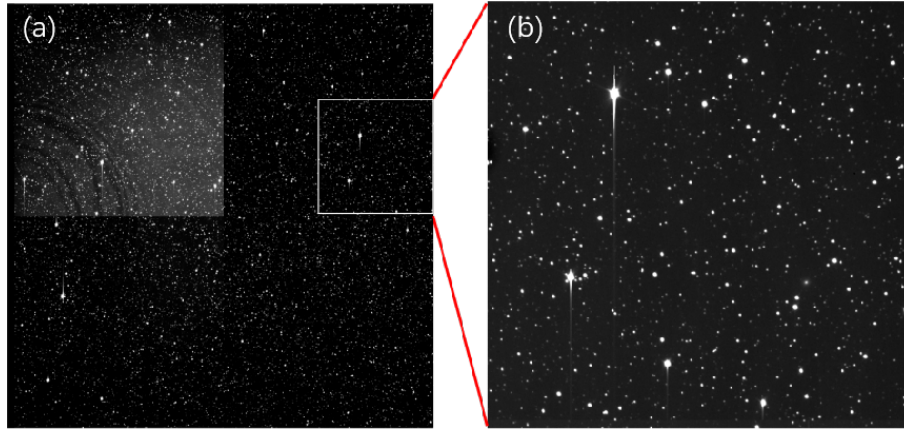


Figure 2. (a) is the initial FITS image and (b) is the FITS image after cropping.

Table 1
Features and Characteristics of the NOWT and its CCD System

Features	Characteristics
Effective diameter	1000 mm
FOV at prime focus	1.5×1.5
Primary mirror reflectance	87%
CCD pixel number	4096×4096
CCD pixel size	$12 \mu\text{m} \times 12 \mu\text{m}$

3.2. WCS Positioning and PINPOINT Astrometry Software

The NOWT used for imaging was mounted in an altitude-azimuth configuration, with its two axes continuously adjusted to track celestial objects as the Earth rotated. The captured images are stored in FITS format, and the corresponding FITS headers contain various metadata, including World Coordinate System (WCS) information (Calabretta & Greisen 2002). Lang et al. (2010) utilized WCS to align images with the celestial sphere coordinate system, enabling accurate positioning of stars and other celestial objects. In the absence of WCS information, we performed a manually assisted asteroid search using THE VISUAL PINPOINT astrometry software. This process involved specifying the image scale and reference catalogs, artificially adding WCS information to the FITS headers, and then manually recording the coordinates of detected asteroids. Within the software, the recognition process involves comparing the image data with catalogs to determine star positions, followed by cross-referencing multiple images to identify moving asteroids. However, we found the software to be cumbersome and time-consuming, with processing times exceeding 20 minutes for only 5–10 images. The result of this operation is shown in Figure 3.

To address these limitations, we developed an independent preprocessing module that employs astronomical image

alignment algorithms to align stellar images. This module superimposes multiple images to reconstruct the asteroid's trajectory, enabling more efficient batch processing and quicker determination of the asteroid's precise position. Specifically, the module automates the alignment and stacking processes, significantly reducing manual intervention and processing time.

3.3. Astronomical Image Alignment Based on Asterism Matching

There are numerous traditional feature-based image registration methods, such as SIFT (Lowe 2004) and ORB (Rublee et al. 2011), most of which rely on identifying local feature points within the image for registration. However, these methods often yield suboptimal results when applied to astronomical images. In such images, both constant stars and asteroids typically appear as point-like light sources, making it challenging for these algorithms to perform accurate recognition and differentiation. To address this issue, Beroiz et al. (2020) proposed a method called Astrolign, which uses asterism matching to establish direct correspondence between two images. This method characterizes star clusters using “geometric hashes,” which are invariant to translation, rotation, and scaling. Consequently, Astrolign is not limited by image source or format and can match images under varying point spread functions and different visual and atmospheric conditions. The primary steps of the method are as follows:

(1) Bright source extraction and triangle construction: Extract the brightest sources from the image, construct all possible triangle combinations, and compute geometric invariants for matching.

(2) Invariant matching and triangle association: Match triangles with similar invariants between two images to establish correspondence between them.

(3) Vertex matching: Determine point-to-point correspondence of triangle vertices based on edge length relationships.

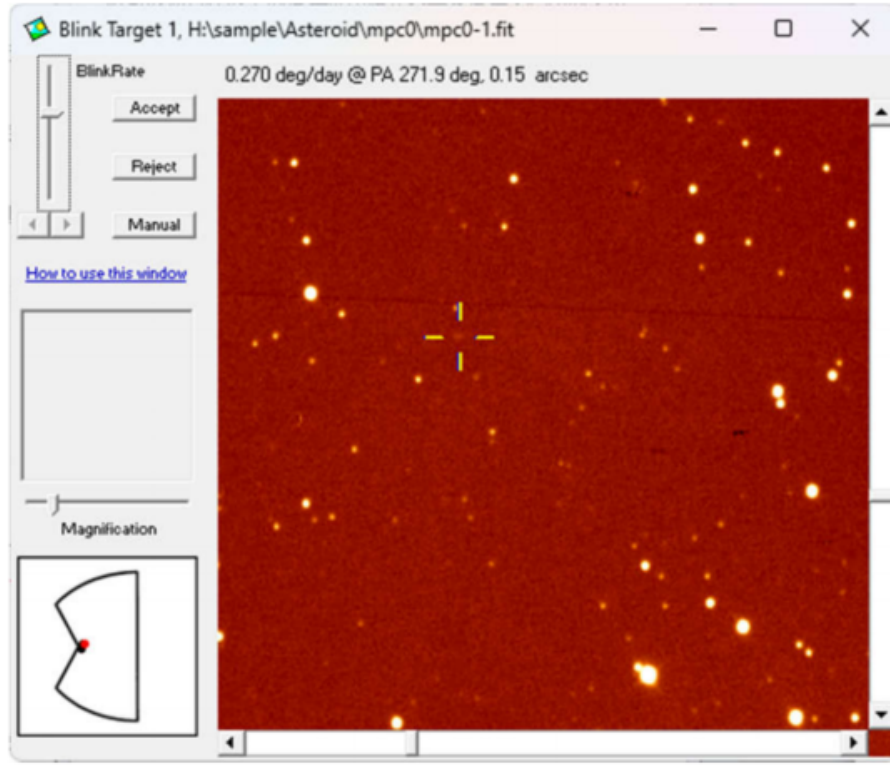


Figure 3. Detection results from the astrometry software Visual Pinpoint.

(4) Transformation fitting: Derive an affine transformation from vertex matches, and accept it if it aligns with most sources in the images.

In short, Astrolign is able to achieve accurate image alignment by finding similar three-point Asterisms (triangles) and deriving affine transformations between them:

$$\mu(\{L_i\}) = \left(\frac{L_2}{L_1}, \frac{L_1}{L_0} \right) \quad (1)$$

Let L_i represent the three sides of a triangle, where $L_2 \geq L_1 \geq L_0$. The mapping μ defines an invariant tuple that fully describes a triangle, including translation, rotation, scaling, and coordinate reflection.

After Astrolign processing of the original image, it is found that the registration between images can be well realized, and the image results after registration are shown in Figure 4.

3.4. Multi-stage Stacking and Greyscale Adjustment for Enhanced Asteroid Detection

Upon completion of the stellar image alignment, the images are divided into unequal time intervals based on the time gaps between the telescope's exposures. The images captured within the same exposure time range are then preliminarily stacked, since the motion of the asteroid is relatively slow and its position changes only marginally over short periods. This

results in the asteroid appearing as a distinct bright spot after the initial stacking, as shown in Figure 5(a)–(b). This process effectively enhances the SNR of the asteroid, thereby improving its contrast. Following the initial stacking, a secondary stacking of the combined images is performed. Although NOWT captures multiple images consecutively during each exposure with minimal variation, the time interval between different exposure periods is relatively long. The intervals between exposures, as indicated in the original FITS images, are 30 minutes each. During the secondary stacking process, the asteroid's unique kinematic properties differentiate it from the background stars, leading to the formation of a distinct trajectory while the background stars remain relatively stationary. This significantly enhances the characterization of the asteroid as a moving object.

During the secondary image stacking process, to prevent an excessive reduction in asteroid brightness, the greyscale values of the five selected images were individually processed. Specifically, for each pixel, the maximum and median greyscale values across the five images were computed, and their average was used as the new greyscale value in subsequent processing. Consequently, the brightness of stationary background stars remained largely unchanged, while the moving asteroid, whose pixel positions varied across the stacked images, had its greyscale values reduced to approximately half of their original brightness by averaging

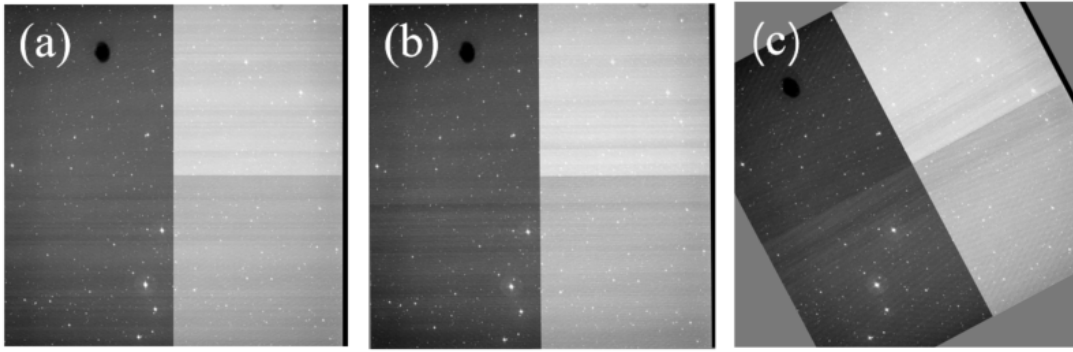


Figure 4. (a) is the baseline image used for registration, (b) is the image after registration with little difference in shooting Angle, and (c) is the image after registration with large difference shooting Angle.

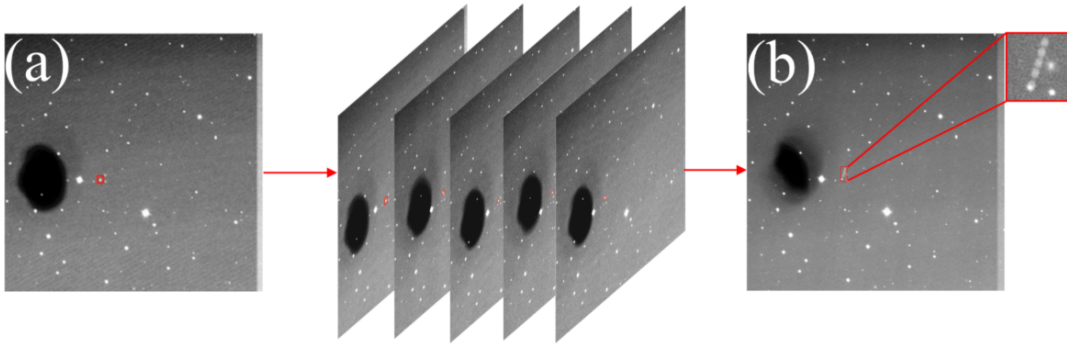


Figure 5. (a) is the result after the first overlay, and (b) is the result after the second overlay of the five images containing the asteroid trajectory object.

the maximum and median values. This approach preserved weak signals that would otherwise be masked by noise and effectively mitigated fluctuations caused by background and transient noise. Compared to directly averaging the greyscale values of the five images, this method better preserved the brightness of the asteroid object, thereby enhancing its contrast and facilitating detection.

3.5. Data Set Composition

We constructed the ATID from the original 16 bit FITS images, which were captured using the NOWT. The images were taken during the observation season from 2015 to 2023, using a target tracking mode with an exposure time of 30 s and a telescope tracking speed of approximately $21''$. The original images ensure high-resolution astronomical data, capturing various celestial objects, including asteroids and stars, across different time intervals. The observations were made in the R band and V band of the Johnson-Cousins $UBVRI$ system (Bessell 1990). After preprocessing the original FITS images, we obtained a total of 6366 images, containing 17,124 asteroid trajectory samples. The 17,124 asteroid trajectory samples in

the data set do not correspond to unique asteroids, as multiple samples may originate from the same asteroid observed across different images or time intervals. The ATID is independently split into three distinct sets: a training set, a validation set, and a testing set, with an approximate ratio of 7:2:1. A detailed description of the ATID is provided in Table 2.

Figure 6 illustrates an example from the constructed asteroid trajectory data set.

3.6. Asteroid Trajectory Confirmation

To verify the reliability of the ATID, we adopted the traditional astronomical processing workflow, that is, we utilized the PinPoint software and the Minor Planet Center (MPC) to verify the corresponding asteroids in the ATID data set. First, we determined the positions of the annotated asteroids in the original images and used the automatic asteroid identification software PinPoint to generate the standardized 80-column MPC format data required by the MPC. Subsequently, these data were submitted to the MPCChecker tool on the MPC for online trajectory matching. During the trajectory matching process, we designated the

Table 2
Asteroid Drag Trajectories Data Set for Astronomical Images

	Number of Images	Number of Samples	Resolution /Pixel
Training set	4456	10,646	1024×1024
Validation Set	1273	3636	1024×1024
Testing set	637	2842	1048×1024

targets with R.A. and decl. deviations less than or equal to $0''.1$ as confirmed matching targets.

By verifying the apparent magnitudes and orbital offsets, we confirmed whether the annotated objects were known asteroids or unknown objects with undetermined asteroid numbers. Whether these objects without asteroid numbers are new asteroid targets requires determination by the MPC. Through this verification process, we identified 350 known asteroids with confirmed asteroid numbers and four unknown objects with undetermined asteroid numbers in the ATID data set. The matching results between ATID's trajectory targets and MPC's real asteroid trajectories can be accessed via the DOI link: <https://doi.org/10.57760/sciencedb.j00167.00023>.

4. Methodology

Object detection is a fundamental task in the field of computer vision. Through object detection algorithms, we can identify objects of interest within an image, accurately locate their positions, and classify them into appropriate categories. Current object detection algorithms based on deep convolutional neural networks can generally be categorized into two-stage and one-stage approaches, depending on the detection process. In two-stage object detection algorithms, the primary objective of the first stage is to generate multiple object candidate regions, while the second stage focuses on filtering these candidate regions, classifying the objects, and refining the predicted locations and sizes. The advantage of this approach lies in the significant improvement in the quality of the predicted regions at each stage, which enhances the overall accuracy of object detection.

To address the issue of low detection rates for dim asteroid objects, we propose an object detection algorithm based on the improved CenterNet2 (Zhou et al. 2021), referred to as AstroCenterNet. The proposed algorithm is a two-phase object detection framework that enhances the detection performance by reducing the FAR and mitigating the susceptibility to noise interference. The first stage of AstroCenterNet employs a ResNet-50 (He et al. 2016) backbone network integrated with the SimAM attention module (Yang et al. 2024), along with an FPN (Lin et al. 2017) that includes a low-level feature fusion branch. The target region proposals are generated using the CenterNet head network. The second stage mirrors the structure of Cascade-RCNN (Cai & Vasconcelos 2018),

refining the object proposals through a cascade of classifiers and bounding box regressors. The overall workflow of the improved CenterNet2 algorithm is illustrated in Figure 7.

4.1. Multi-channel Histogram Truncation Module

Astronomical greyscale data are typically represented in 16 bit format, with pixel values ranging from 0 to 65,536. However, the brightness of astronomical objects often falls within a relatively narrow range. Direct normalization of these values may cause dim objects, such as asteroids, to be obscured by the background. To address this issue, we used a preprocessing module known as the Multi-channel Histogram Truncation Module (MHT; Jiang et al. 2023). The MHT processing steps are illustrated in Figure 8.

First, the MHT module performs histogram analysis on the input image and applies a smoothing operation to the resulting histogram curve. As shown in Figure 8, the effective greyscale values of astronomical images are concentrated within a narrow range. A threshold is applied to the number of pixels, and any pixels below this threshold are assigned a value of zero. Within the truncation range, the greyscale values are further divided into high, medium, and low channels, which are normalized separately. These channels are then concatenated along the channel dimension and passed into the subsequent network for further processing. We preprocess the 16 bit astronomical images using MHT, which significantly enhances the features of the objects. In the first row of Figure 9, the object is barely visible due to its dimness; however, after preprocessing, the object becomes clearly discernible in the low and high channels. In the second row of Figure 9, the object is almost invisible to the naked eyes, but the features of the object are greatly enhanced in the high-channel image obtained after MHT preprocessing.

4.2. SimAM Attention Module

Since small objects are susceptible to noise interference, we hope that neural networks can pay more attention to pixel regions with greater background differences. Therefore, this study proposes using the SimAM attention mechanism to optimize the backbone network and improve the network's ability to extract features from small objects.

SimAM (Yang et al. 2024) is a parameter-free attention mechanism proposed by Sun Yat-sen University in 2021. Based on the well-known neuroscience theory, it optimizes the energy function to find the importance of each neuron, and further simplifies the module by solving the analytical solution of the energy function. Another advantage of the SimAM module is that most operations are based on direct calculations of the defined energy function, so there is no need for excessive network structure adjustments, which is flexible and effective in improving the representation ability of convolutional networks.

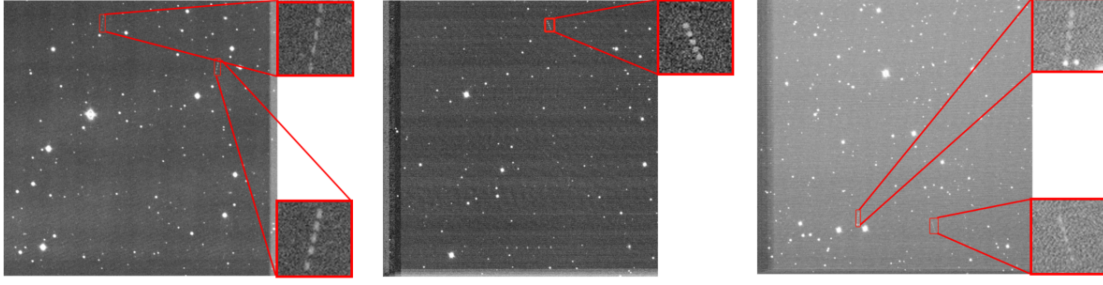


Figure 6. Example images of a self-developed data set of asteroid drag trajectories.

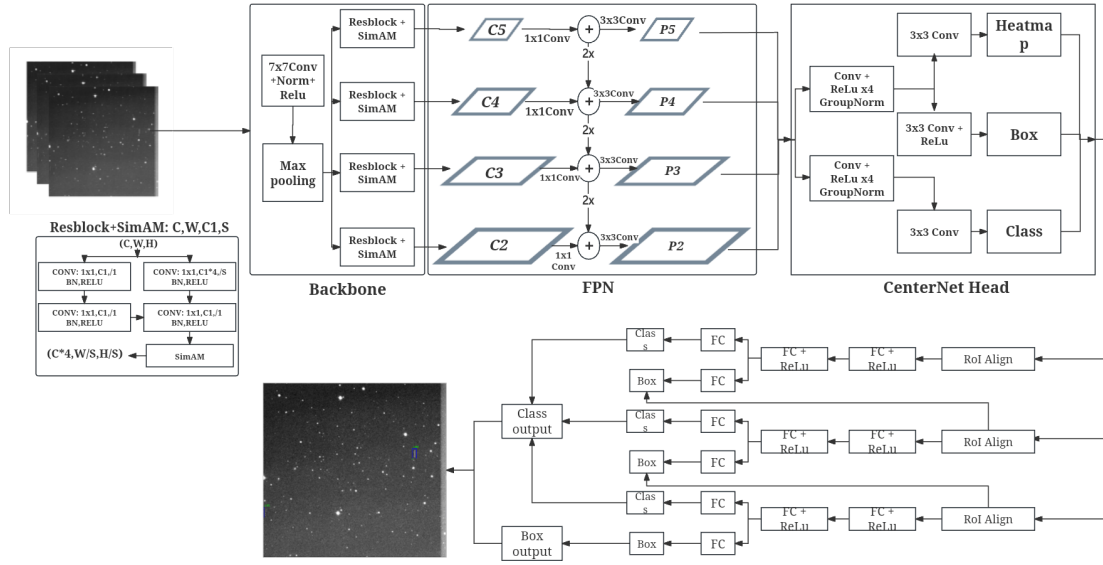


Figure 7. The network architecture of AstroCenterNet.

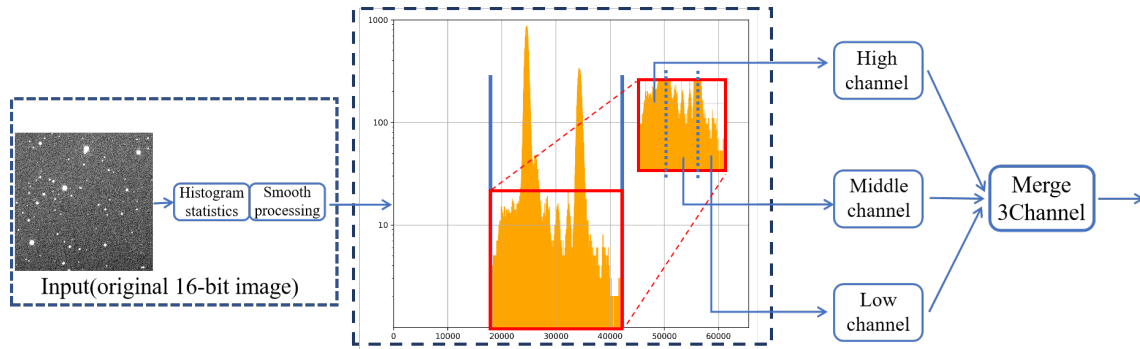


Figure 8. Processing Flow of MHT.

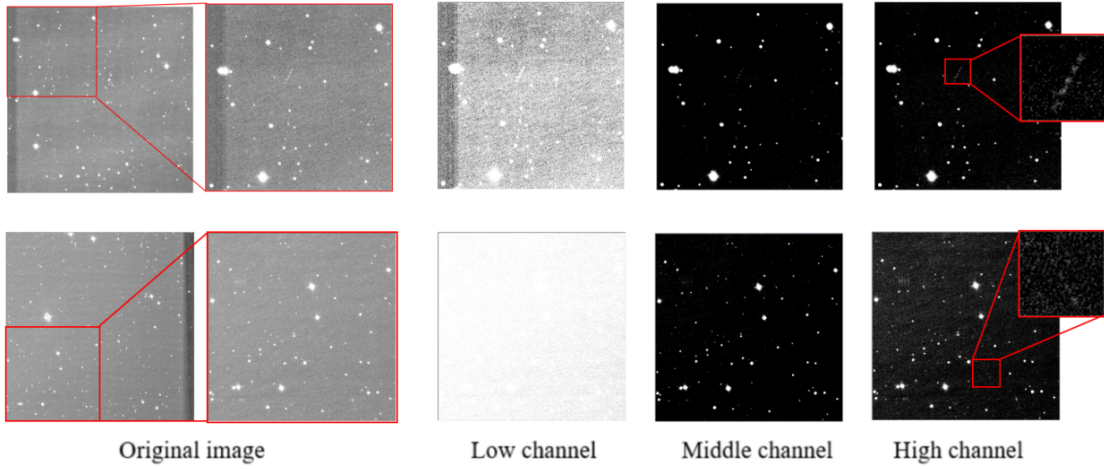


Figure 9. Processing effect of MHT: (a) original image, (b) low channel, (c) middle channel, and (d) high channel.

Adding neural networks can make the network more focused on those pixel points in the input image that have large differences from surrounding pixels. The minimum energy can be obtained by the following Equation (2):

$$e_t^* = \frac{4(\sigma^2 + \lambda)}{(t - u)^2 + 2\sigma^2 + 2\lambda} \quad (2)$$

where t denotes the value of each pixel in the image, u denotes the pixel mean of the image, σ denotes the channel variance, and λ is the manually set hyperparameter.

The lower the energy, the greater the difference between the neuron and its surrounding neurons, and the higher its importance. Therefore, the importance of a neuron can be obtained by taking its reciprocal value. According to the definition of attention mechanism, the minimum energy matrix is finally multiplied by the input feature map after being passed through the sigmoid activation function. By calculating the above energy function, the SimAM attention mechanism obtains the 3D weights of the current feature map and assigns a unique weight to each neuron.

SimAM essentially calculates the importance of each element in the 3D feature map for both spatial and channel dimensions simultaneously, in order to highlight the most important features. Finally, this study uses the SimAM attention mechanism to improve the residual block structure of the backbone network.

4.3. Low-level Feature Detection Branch

Neural networks usually improve their grasp of object-semantic information by increasing the depth of the network. As the network gradually reduces the sample size, the spatial resolution gradually decreases and the receptive field of pixels gradually increases, making the high-dimensional semantic information of the target more obvious. This plays an important role in improving

the performance of various computer vision tasks such as image classification and object detection. However, in more challenging tasks with low image resolution or small objects, their performance rapidly declines. For small objects, as the network downsamples, features are more likely to be lost. Therefore, researchers use feature fusion modules to fuse feature maps of different scales (Quan et al. 2023; Zhang et al. 2023; Chen et al. 2024), such as FPNs (Lin et al. 2017), to improve the network's ability to recognize objects of different sizes.

CenterNet2 uses traditional FPNs as the feature fusion module to extract deeper feature maps for fusion. Although it may contain more background noise and bring more computation, we still claim that the Low-level feature map is equally important for improving small objects. Therefore, based on the original three FPN detection branches, we added a low-level feature map detection branch to improve the detection ability of the network for small objects, as shown in Figure 10.

5. Results and Discussions

5.1. Experimental Details

The algorithms presented in this paper are implemented using the Python programming language. Network training is carried out on a laboratory server running Ubuntu 20.04.1 as the operating system. The network architecture is constructed using the PyTorch framework, with the NVIDIA CUDA library employed for deep learning acceleration. The server hardware configuration includes eight NVIDIA RTX A5000 graphics cards, each equipped with 24 GB of video memory, 512 GB of RAM, and two 10-core, 16-thread CPUs operating at a base clock speed of 2.90 GHz. Detailed hardware specifications are provided in Table 3.

In addition, uniform hyperparameter settings are applied during the model-training process. The batch size is set to 16,

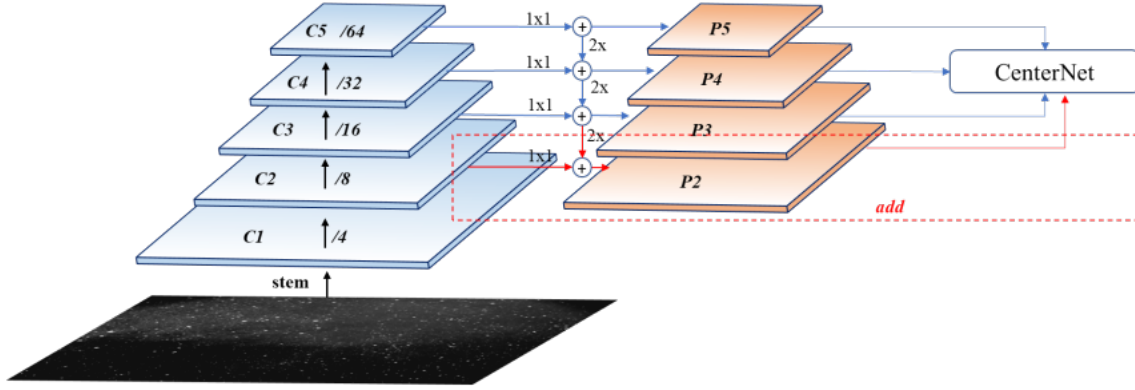


Figure 10. Improved FPN module, adding a low-resolution feature fusion and low-resolution feature map detection module.

the learning rate is 0.01, and the number of training epochs is set to 200. For model evaluation, the Intersection over Union (IoU) threshold for determining the success of detection is set to 0.5, which allows for a better assessment of the model's capability to accurately locate small objects.

5.2. Evaluation Metrics

This section outlines the evaluation metrics employed to assess the performance of the object detection network. In this study, Precision, Recall, F_1 Score, Average Precision (AP) (Everingham et al. 2010), and FAR are utilized as primary metrics, with their respective formulas provided in Equations (3), (4), (5), (6) and (7).

True Positive (TP): Refers to the positive samples predicted as positive by the model. The output value of the detection model is the predicted value, and each sample within the prediction box is considered as a positive prediction. TP is defined when the IoU between the prediction box and the annotated box exceeds a certain threshold (commonly set at 0.5).

True Negative (TN): Refers to the negative samples predicted as negative by the model.

False Positive (FP): Refers to the negative samples predicted as positive by the model, also known as a false alarm.

False Negative (FN): Refers to the positive samples predicted as negative by the model, also known as a missed detection.

Precision (P): The proportion of all samples predicted by the model to be in the positive category that are actually in the positive category.

$$P = \frac{TP}{TP + FP} \quad (3)$$

Recall (R): The ratio of the number of correctly detected samples to the total number of samples.

$$R = \frac{TP}{TP + FN} \quad (4)$$

Table 3
Asteroid Drag Trajectories Data Set for Astronomical Images

Hardware	Server	Amax G448-X2
	CPU	Intel(R)Xeon(R)Gold CPU @2.90 GHz x2
	GPU	RTX A5000 x8
	Memory	64 GB x8
Software	Operating system	Ubuntu20.04.1
	Compiler	Pycharm 2020.3
	Programming language	Python 3.8
	Acceleration library	CUDA 11.1+cu111
	Library for Network Building	Torch1.8.1

F_1 Score: An approach that combines model precision and recall.

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (5)$$

Average Precision (AP): Under the precision-recall curve, the area under the curve is calculated. This area is the AP value, which represents the combined performance of the model under different recall thresholds.

$$AP = \sum_n P(n) \Delta R(n) \quad (6)$$

False Alarm Rate (FAR): The ratio of the total number of false alarms to the total number of predictions.

$$FAR = \frac{FP}{TP + FP} \quad (7)$$

5.3. Experimental Results

The training of neural networks relies on data sets with annotated ground truth, and the composition of the data set, the quality of annotations, and the volume of data have a direct impact on the training process and the final performance of the model. In this study, the self-constructed asteroid trajectory

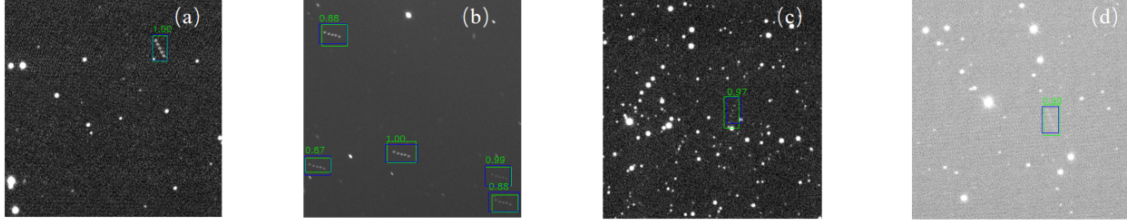


Figure 11. Cropped display of some of the detections in the testing set, with the numbers representing the confidence level of the detections: (a) a display of the most prevalent asteroid detections, (b) a display of multiple asteroid object detections, (c) a display of faint asteroid object detections in a dense background object, and (d) a display of detections of cloud-obscured faint object.

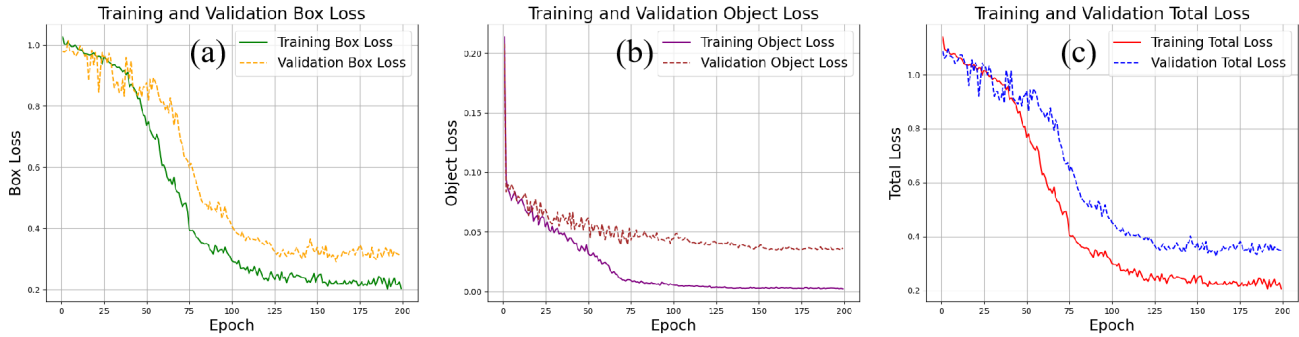


Figure 12. Total Loss = box Loss + object Loss + cls Loss, with cls Loss = 0 due to a single classification task.

astronomical image data set is randomly split into three subsets: a training set, a validation set, and a test set, with a ratio of 7:2:1, as detailed in Table 2. The validation set is used to monitor the convergence of the model during training, while the test set is employed for final evaluation and fine-tuning to ensure that the model maintains accuracy despite the presence of complex background information and noise interference.

In object detection, the network generates multiple prediction frames, each associated with a confidence score. Initially, the prediction frames are filtered by comparing their confidence scores against a predefined threshold. Frames with confidence scores above the threshold are considered positive detections, while those below the threshold are discarded.

The training process consists of 200 epochs, with an initial learning rate of 0.01 and a batch size of 16. Upon completion of the training, the final model is evaluated on the test set. The results show that AstroCenterNet achieves a precision of 98.4%, a recall of 97.6%, and an average precision AP of 94.0% at an IoU threshold of 0.5. Additionally, the false alarm rate is only 1.6%. Some of the detection results are shown in Figure 11.

In object detection, the total loss is the aggregate of all loss terms, which are used to evaluate the model's overall performance. The objective is to minimize this total loss in

order to optimize the model. Box loss (Girshick et al. 2014) measures the positional discrepancy between the predicted bounding box and the ground truth box. It typically quantifies the difference using translation deviation or IoU to assess the accuracy of the predicted box. Object loss (Liu et al. 2016) is a loss function designed to evaluate the model's performance in recognizing objects within the predicted bounding box. This term is primarily used to determine the presence of objects within the predicted box and to assess the correctness of the predicted object category. The loss curve throughout the training process is depicted in Figure 12.

All asteroid trajectory objects in the test set were annotated, and their corresponding original FITS images were retrieved. The SNR of each object within these images was evaluated using the astronomical software MAXLM DL. The SNR for each synthesized image was computed by averaging the SNR values across the individual images, with the resulting average serving as the SNR for the corresponding object. The outcomes of this analysis are presented in Figure 13. Statistical analysis of the SNR for missed detections reveals that the SNR of undetected objects falls within the range of 2–15, with a concentration observed between 2 and 10. For the purpose of observation, an asteroid target with a detectable SNR value

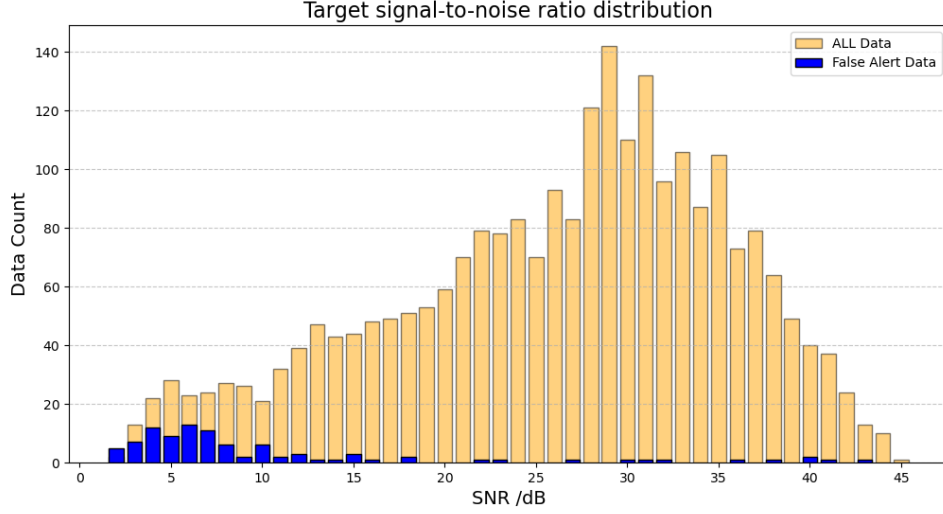


Figure 13. Distribution of signal-to-noise ratio of the testing set.

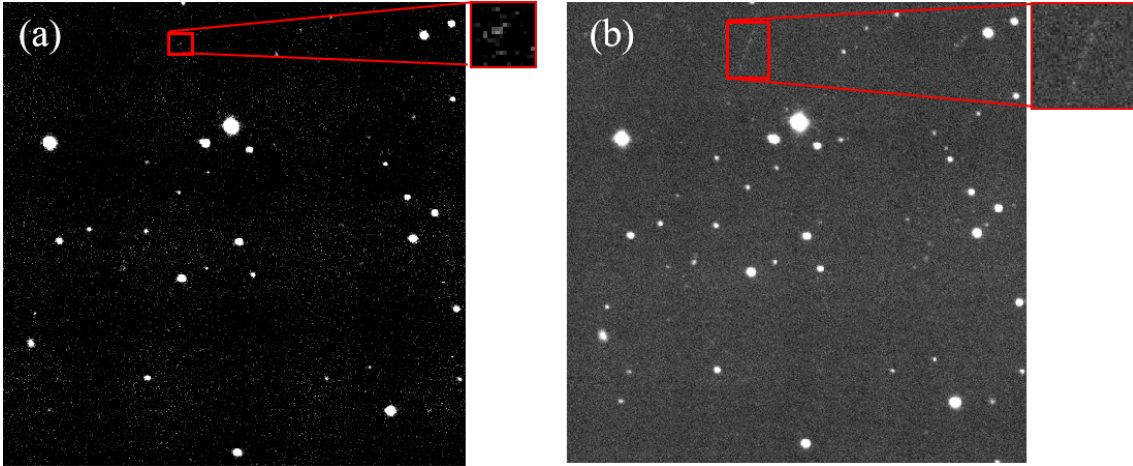


Figure 14. The object signal-to-noise ratio is around 2 for asteroid objects. (a) is the original image and (b) is the image after processing.

close to 2 decibels was selected from the detection results, as shown in Figure 14, where panel (a) displays the original image and panel (b) presents the image after processing. As the SNR decreases, the likelihood of false positives increases. Notably, when the SNR is approximately 5, the false alarm rate escalates significantly. Conversely, when the SNR exceeds 10, the object recall rate surpasses 98%.

By analyzing the detection results on the test set, several types of misidentified objects were observed: (a) In astronomical images, it appears as striped objects, which suspected to be an artificial satellite or space debris (Virtanen et al. 2016) or space debris (Xi et al. 2020); (b) The brightness value is significantly higher than the surrounding pixels, mainly cosmic rays (Wang et al. 2024); and (c) Suspected offset error or noise

artifact (Meredith et al. 2022; Cowan et al. 2023), as illustrated in Figure 15.

Through our further analysis, we speculate that the fundamental causes of these false alarms mainly include insufficient training data (lack of diverse non—asteroid samples), the limitations of feature extraction, and preprocessing errors (such as noise and artifacts). To further improve the detection performance, we plan to conduct comprehensive verification of these false alarm targets in the future, exploring their true corresponding properties of celestial bodies. To this end, we will carry out precise astronomical positioning and photometric calculations on these targets in the next step of our work. By extracting key information from the targets, we will convert suspicious misidentification into confirmed

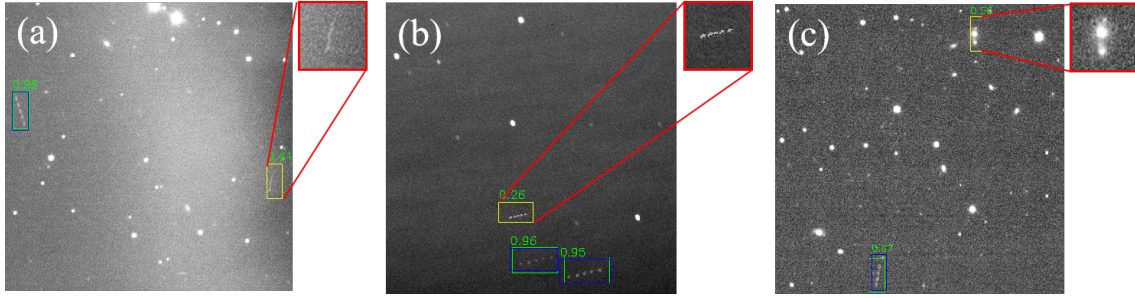


Figure 15. Misidentified objects: (a) Striped object, suspected to be an artificial satellite or space debris; (b) Large brightness difference, suspected to be cosmic rays; (c) Suspected offset error or noise artifact.

Table 4
Comparison of Different Object Detection Algorithms for the Same Hyperparameter on the Same Device

Algorithm	Precision (%)	Recall (%)	AP (%)	FPS (s^{-1})	F1-score (%)	FAR (%)
YOLOv3 (Redmon 2018)	85.4	71.2	80.4	5.6	77.7	14.6
YOLOv5	84.9	75.7	82.6	20.3	80.0	15.1
YOLOv7 (Wang et al. 2023)	86.2	78.2	84.4	13.7	82.0	12.8
YOLOv8	88.2	78.2	84.4	87.7	82.9	11.8
YOLOv11	90.6	80.5	86.9	16.7	85.3	9.4
Faster R-CNN (Ren et al. 2016)	92.9	94.9	95.4	11.9	93.9	7.1
CenterNet (Duan et al. 2019)	87.3	90.0	90.8	16.4	88.6	12.7
CenterNet2	92.2	91.5	89.8	15.6	91.9	7.8
AstroCenterNet(Ours)	98.4	97.6	94.0	17.8	98.0	1.6

Note: Bold values indicate the optimal result for each evaluation metric across all compared experiments.

recognition types. Based on these confirmed results, we will further optimize the training data set to enhance the detection network's ability to recognize such false alarm targets, ultimately improving detection accuracy.

5.4. Comparative Experiment and Ablation Experiment

Through an in-depth study of the dark, weak, and tiny characteristics of space object imaging in large field-of-view astronomical images, this paper proposes several improvements to the CenterNet2 algorithm, resulting in the development of AstroCenterNet. We conducted a comparative analysis of the detection metrics for classical object detection algorithms across different periods on the data set, including YOLOv3 (Redmon 2018), YOLOv5, YOLOv7 (Wang et al. 2023), and CenterNet (Duan et al. 2019). The results are summarized in Table 4. It is evident that AstroCenterNet outperforms other current object detection algorithms in terms of precision and recall. Additionally, we compare the P-R curve of the original network CenterNet2 with that of our improved network AstroCenterNet at an IoU threshold of 0.5, as shown in Figure 16. The comparison reveals that the AP of our improved AstroCenterNet network is significantly higher than that of the original CenterNet2 network. To evaluate the effectiveness of

these improvements, we conducted ablation experiments to assess the impact of each module on detection accuracy. The results, presented in Table 5, indicate that the improved modules enhance performance on the data set, not only improving precision and recall but also reducing the false alarm rate.

6. Conclusions

This study proposes an integrated framework for asteroid detection in large-field astronomical images, which combines data preprocessing and deep learning optimization to address the challenges of low SNR, background noise interference, and weak target recognition in asteroid detection. In long-exposure observations with wide-field telescopes, asteroids exhibit distinct kinematic characteristics due to their significantly different motion speeds compared to background stars. By leveraging stellar cluster registration and multi-frame fusion, the framework extracts temporal features and enhances dynamic features, transforming asteroid objects into discontinuous moving trajectory objects with unique characteristics, thereby improving the detection capability for weak targets and significantly enhancing the separation between targets and background in complex scenes.

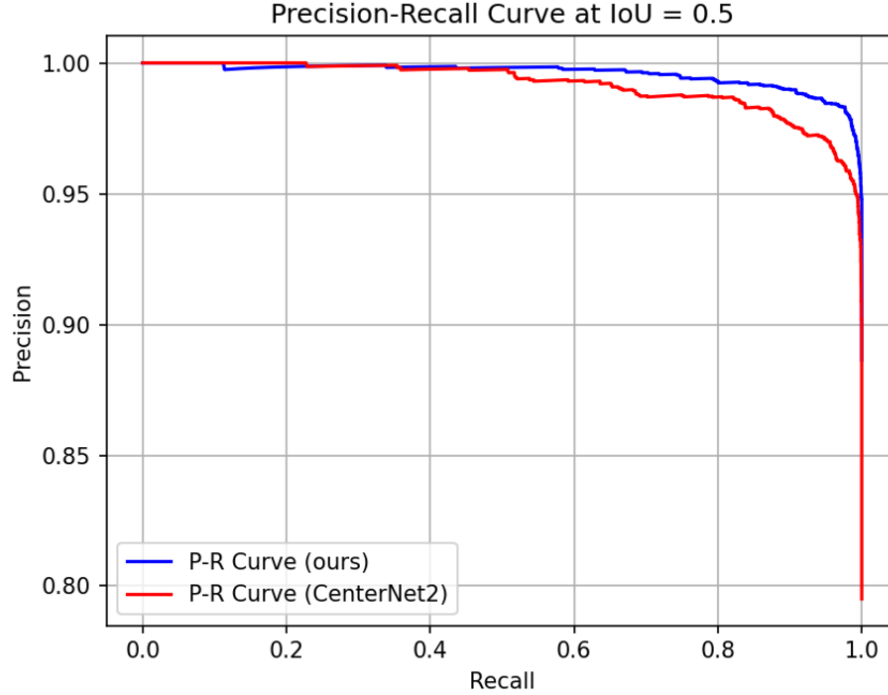


Figure 16. The precision-recall curves for the original CenterNet2 model and our AstroCenterNet model at IoU = 0.5.

Table 5
CenterNet2 Algorithm for Ablation Experiments

Algorithm	Precision % (%)	Recall (%)	AP (%)	FPS (s ⁻¹)	F1-score (%)	FAR (%)
CenterNet2	92.2	91.5	89.86	15.6	91.9	7.8
CenterNet2+MHT	93.5	94.9	91.11	20.3	94.2	6.5
CenterNet2 +MHT +SimAM	94.7	96.3	92.06	13.7	95.5	5.3
CenterNet2+MHT +SimAM+ Low feature detection	98.4	97.6	94.01	17.8	98.00	1.6

Note: Bold values indicate the optimal result for each evaluation metric across all compared experiments.

Furthermore, we have proposed specific improvements for deep learning frameworks, such as the integration of MHT, SimAM attention mechanism, and an improved FPN. The combination of these optimized modules enhances the capability of deep learning models to identify asteroid objects in wide-field astronomical images. Experimental results demonstrate that this method significantly improves asteroid detection accuracy in complex observational environments while effectively reducing false positives and missed detections.

Beyond its applicability in asteroid detection, this method can also be extended to other astronomical target detection tasks with low SNR, such as binary star identification, variable star monitoring, and automatic recognition of other transient stars. Future research could explore cross-band data fusion to enhance detection robustness under varying observational conditions. Additionally, optimizing computational efficiency remains a critical direction, such as employing lightweight

models to reduce computational resource demands and enhance real-time detection capabilities. These advancements will contribute to large-scale automated astronomical target detection and support the development of efficient and intelligent astronomical observation systems.

Acknowledgments

This research was funded by the National Science and Technology Major Project (2022ZD0117401) and the National Defense Science and Technology Innovation Special Zone Project Foundation of China (grant No. 19-163-21-TS-001-067-01). In addition, support was provided by the Chinese Academy of Sciences (CAS) “Light of West China” Program (No. 2020-XBQNXZ-016 and No.2022-XBQNXZ-016). We would also like to thank the Nanshan One-meter Wide-field Telescope (NOWT) (IAU code N87) of Xinjiang Astronomical Observatory for providing us with the original astronomical image data

Data Availability

The data will be shared on reasonable request to the corresponding author. The software utilized in this research is proprietary and is not openly accessible to the public. For any collaboration initiatives or inquiries specifically pertaining to the software, kindly address your correspondence to the corresponding author, who will undertake an assessment of the viability of software sharing contingent upon reasonable and mutually beneficial proposals.

References

- Azadmanesh, M., Roshanian, J., & Hassanalian, M. 2023, *PrAeS*, **142**, 100957
- Bai, C., Feng, G., Zhang, X., et al. 2020, *RAA*, **20**, 211
- Beroiz, M., Cabral, J., & Sanchez, B. 2020, *A&C*, **32**, 100384
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Bessell, M. S. 1990, *PASP*, **102**, 1181
- Bochkovskiy, A., Wang, C., & Liao, H. 2020, arXiv:2004.10934
- Cai, Z., & Vasconcelos, N. 2018, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR), ed. B. Michael (Piscataway, NJ: IEEE), 6154
- Calabretta, M. R., & Greisen, E. W. 2002, *A&A*, **395**, 1077
- Chen, Y., Zhang, C., Chen, B., et al. 2024, *Comput. Biol. Med.*, **170**, 107917
- Cowan, P., Bond, I. A., & Reyes, N. H. 2023, *A&C*, **42**, 100693
- Dai, Y., Zheng, T., Xue, C., & Zhou, L. 2022, *ApSci*, **12**, 4610
- Day, P., Crane, J., Cronk, P., et al. 2024, in Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conf., ed. S. Ryan (Maui, Hawaii: MEDB), 54
- Duan, K., Bai, S., Xie, L., et al. 2019, in Proc. IEEE/CVF Int. Conf. Comput. Vis., ed. D. Larry (Piscataway, NJ: IEEE), 6569
- Du, Z., Jiang, H., Yang, X., et al. 2024, *AdSpR*, **73**, 5349
- Everingham, M., Van Gool, L., Williams, C. K. I., et al. 2010, *Int. J. Comput. Vis.*, **88**, 303
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. 2014, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., ed. B. Ronen (Piscataway, NJ: IEEE), 580
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR), ed. L. Agapito (Las Vegas, NV: IEEE), 770
- He, Z., Qiu, B., Luo, A., et al. 2021, *MNRAS*, **508**, 2039
- Hobson, M. P., & McLachlan, C. 2003, *MNRAS*, **338**, 765
- Huang, X., Wang, X., Lv, W., et al. 2021, arXiv:2104.10419
- Jiang, P., Liu, C., Kang, Z., Yang, W., & Li, Z. 2022, *RAA*, **22**, 105003
- Jiang, Y., Tang, Y., & Ying, C. 2023, *Electronics*, **12**, 4820
- Lang, D., Hogg, D. W., Mierle, K., et al. 2010, *AJ*, **139**, 1782
- Lin, T., Dollar, P., Girshick, R., et al. 2017, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR), ed. C. Rama (Piscataway, NJ: IEEE), 936
- Liu, D., Chen, B., Chin, T., & Rutten, M. G. 2020a, *ITSP*, **68**, 5166
- Liu, D., Wang, X., Xu, Z., et al. 2020b, *A&C*, **32**, 100408
- Liu, W., Anguelov, D., Erhan, D., et al. 2016, *Comput. Vis.—ECCV 2016*, 9905, 21
- Lowe, D. G. 2004, *Int. Journ. of Computer Visi.*, **60**, 91
- Masias, M., Freixenet, J., Lladó, X., & Peracaula, M. 2012, *MNRAS*, **422**, 1674
- Meredith, C., Privett, G., George, S., & Feline, W. 2022, Novel image alignment technique for extraction of astrometry and photometry from small field of view astronomical sensors, Adv. Maui Opt. Space Surveill. Technol. Conf., ed. S. Ryan,
- Ohsawa, R. 2022, *Journal of Space Science Informatics Japan*, **11**, 1
- Pan, H., Song, G., Xie, L., & Zhao, Y. 2014, *OptRv*, **21**, 292
- Pan, Z., Qiu, B., & Li, G. 2024, *MNRAS*, **534**, 726
- Petit, J., Holman, M., Scholl, H., Kavelaars, J., & Gladman, B. 2004, *MNRAS*, **347**, 471
- Pontinen, M., Granvik, M., Nucita, A. A., et al. 2023, *A&A*, **679**, A135
- Pradhan, B., Hickson, P., & Surdej, J. 2019, *AcAau*, **164**, 77
- Quan, Y., Zhang, D., Zhang, L., & Tang, J. 2023, *ITIP*, **32**, 4341
- Redmon, J. 2018, arXiv:1804.02767
- Ren, S., He, K., Girshick, R., & Sun, J. 2016, *ITPAM*, **39**, 1137
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. 2011, in Int. Conf. Comput. Vis., ed. A. Amit (Piscataway, NJ: IEEE), 2564
- Sun, R., Yu, S., Zhao, C., & Zhang, W. 2019, *PASJ*, **71**, 67
- Sun, R., Zhan, J., Zhao, C., & Zhang, X. 2015, *AcAau*, **110**, 9
- Sun, R., & Zhao, C. 2013, *RAA*, **13**, 604
- Virtanen, J., Poikonen, J., Sääntti, T., et al. 2016, *ASR*, **57**, 1607
- Wang, C., Bochkovskiy, A., & Liao, H. 2023, in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., ed. M. Brown (Piscataway, NJ: IEEE), 7464
- Wang, L., Zhang, X., Bai, C., et al. 2024, *MNRAS*, **534**, 385
- Xi, J., Xiang, Y., Ersoy, O. K., et al. 2020, *IEEEA*, **8**, 150864
- Yang, L., Zhang, R., Li, L., & Xie, X. 2024, in Int. Conf. Mach. Learn., ed. M. Meila (PMLR), 110509
- Zhang, G., Li, Z., Li, J., & Hu, X. 2023, arXiv:2302.06052
- Zhou, X., Koltun, V., & Krahenbuhl, P. 2021, arXiv:2103.07461