ChinaRxiv 中X

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202412.00103

Deep Neural Network Closed-loop with Raw Data for Optical Resident Space Object Detection postprint

Authors: He Zhao, Rong-Yu Sun and Sheng-Xian Yu

Date: 2024-12-06T11:39:37+00:00

Abstract

Optical survey is an important means for observing resident space objects and space situational awareness. With the application of astronomical techniques and reduction method, wide field of view telescopes have made significant contributions in discovering and identifying resident space objects. However, with the development of modern optical and electronic technology, the detection limit of instruments and infrastructure has been greatly extended, leading to an extensive number of raw images and many more sources in these images. Challenges arise when reducing these data in terms of traditional measurement and calibration. Based on the amount of data, it is particularly feasible and reliable to apply machine learning algorithms. Here an end-to-end deep learning framework is developed, it is trained with a priori information on raw detections and the automatic detection task is performed on the new data acquired. The closed-loop is evaluated based on consecutive CCD images obtained with a dedicated space debris survey telescope. It is demonstrated that our framework can achieve high performance compared with the traditional method, and with data fusion, the efficiency of the system can be improved without changing hardware or deploying new devices. The technique deserves a wider application in many fields of observational astronomy.

Full Text

Preamble

Research in Astronomy and Astrophysics, 24:115009 (9pp), 2024 November
© 2024. National Astronomical Observatories, CAS and IOP Publishing Ltd. All rights reserved, including for text and data mining, AI training, and

ChinaRxiv 中X

similar technologies. Printed in China.
https://doi.org/10.1088/1674-4527/ad849a

Deep Neural Network Closed-loop with Raw Data for Optical Resident Space Object Detection

He Zhao[1,2], Rong-Yu Sun[1], and Sheng-Xian Yu[1]

[1] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China; rysun@pmo.ac.cn
[2] School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China

Received 2024 June 13; revised 2024 August 25; accepted 2024 October 5; published 2024 November 4

Abstract

Optical surveys represent an important means for observing resident space objects (RSOs) and supporting space situational awareness. With the application of astronomical techniques and reduction methods, wide-field-of-view telescopes have made significant contributions to discovering and identifying RSOs. However, modern optical and electronic technologies have greatly extended the detection limits of instruments and infrastructure, leading to an explosion in raw image volume and source density. These developments create substantial challenges for traditional measurement and calibration pipelines when reducing such data. Given the sheer volume of data, machine learning algorithms become particularly feasible and reliable. Here we develop an end-to-end deep learning framework trained with a priori information from raw detections to perform automatic detection on newly acquired data. We evaluate this closed-loop system using consecutive CCD images from a dedicated space debris survey telescope. Our framework achieves high performance compared to traditional methods, and through data fusion, system efficiency can be improved without hardware modifications or new device deployments. This technique merits wider application across many fields of observational astronomy.

Key words: methods: data analysis – techniques: image processing – surveys

1. Introduction

The development of satellite and launch technologies has led to a significant proliferation of resident space objects in orbit, including communication and navigation satellites, rocket bodies, and consequent space debris, all of which threaten the safety and sustainability of the space environment. Obtaining effective information on RSO positions and status—known as space situational awareness (SSA)—is therefore essential and crucial. Optical surveys using astronomical telescopes represent an important passive means for observing and detecting RSOs (Schildknecht 2007). Compared to active methods like laser ranging (Zhang et al. 2012) and radar, optical surveys are more appropriate for high-Earth orbital regions (Matney et al. 2004; Sun et al. 2015). Furthermore, considering the relatively low economic cost of optical infrastructure, it is feasible to develop telescope arrays (Zhang & Zhao 2021) or networks with multiple sites (Molotov et al. 2008; Tingay et al. 2013).

Wide-field-of-view telescopes are widely used in optical RSO surveys because they can survey large sky areas in shorter times, thereby promoting efficiency. Consequent object detection and extraction algorithms have been developed and deployed, among which traditional astronomical source extraction techniques such as SExtractor (Bertin & Arnouts 1996) and DAOPHOT (Stetson 1987; Schechter et al. 1993) are widely applied and have played important roles. Meanwhile, dedicated algorithms have also been proposed for specific applications and achieve excellent performance in object detection.

In detail, considering the relative movement between background stars and RSOs, the images of stars and RSOs appear with different shapes depending on observing strategy. Techniques utilizing masking or streak detection work effectively in data reduction (Kouprianov 2008; Sun et al. 2016; Hickson 2018). Based on the movement characteristics of RSOs, methods have also been proposed to extract and correlate objects from consecutive frames in celestial coordinates (Sun et al. 2019; Du et al. 2022; Zhang et al. 2024). Furthermore, image processing methods including morphology transformation (Sun & Zhao 2013) and restoration (Sun & Jia 2017), as well as image stacking (Yanagisawa et al. 2005), are widely used to improve the signal-to-noise ratio (SNR) of objects and promote detection efficiency.

However, with the development of telescope and sensor technology, the detection limit of modern infrastructure has been greatly extended, leading to an explosion in raw image volume and source density. For example, the number of sources in a single wide-field image can exceed 10,000, and due to the deployment of sensors with fast readout speeds, frame rates have increased significantly, yielding extensive data volumes. These developments create challenges for previous data reduction and object detection methods, affecting extraction efficiency and time cost while limiting overall performance.

The development of artificial intelligence has triggered a technological revolution, producing various object detection models that, after training, can extract different source types from single frames akin to human perception. These techniques require large amounts of data to train model parameters for specific tasks, making them particularly feasible given the big data obtained by modern astronomical infrastructure. Significant breakthroughs have been made in these areas. For the Sloan Digital Sky Survey (SDSS), YOLOv4 has been adopted to develop a source detection and classification network (He et al. 2021). A modified YOLOv3 has been presented for redshift galaxy cluster detection, achieving performance comparable to traditional methods (Grishin et al. 2023). Convolutional neural networks (CNNs) are also widely applied in data reduction for wide-field telescopes (Jia et al. 2020), such as detecting images with asteroid streaks (Wang et al. 2022a) and finding blue horizontal-branch stars (He et al. 2023). In addition to these one-stage methods, two-stage-based models are also used for point and streak source detection in dedicated applications (Dumitrescu et al. 2022).

However, it should be noted that in these applications, models are trained and validated mostly on simulated data. Considering the lack of raw training data, performance remains limited and deserves further investigation.

In our work, we develop a deep learning neural network based on the YOLOv5 model to detect and extract RSOs from large volumes of raw CCD images. Our method is trained on raw data with a priori information obtained through traditional methods. Based on morphological differences between stars and RSOs, the model learns these features and extracts RSO images from single frames along with their measurement information. We evaluate performance on raw data and demonstrate that our closed-loop network achieves high efficiency. With data fusion, system performance can be improved. In Section 2 we introduce the principles and algorithms, Section 3 describes the application, Section 4 discusses the results, and Section 5 presents our conclusions.

2. Principles and Algorithms

The deep learning process includes two key steps: dataset construction and neural network architecture. In this era of big data, high-quality data play a crucial role in deep learning, with most datasets constructed through manual labeling and annotation. Meanwhile, an appropriate architecture exhibits excellent performance for specific tasks. For dataset construction, we propose an automatic labeling method that utilizes previously applied object detection algorithms to extract RSO image positions in real time, building the training dataset from this information. This approach saves extensive time and generates a comprehensive dataset. Our method also optimizes label parameters based on distributions and experimental results.

Furthermore, we employ image transformations with a priori knowledge to promote object detection and improve efficiency. For the network architecture, considering that RSO images appear as small-scale and faint sources in large field-of-view frames, we adjust the network structure in the neck and head areas to better detect objects in these specific categories.

2.1. Network Architecture

YOLO represents a cutting-edge convolutional neural network (CNN) that provides real-time object detection capability. Unlike two-stage CNN detection networks that utilize separate networks to first detect candidates and then classify them, YOLO employs a single network for simultaneous positioning and classification, achieving end-to-end detection that fulfills our requirements well. During reduction, it divides input data into grid cells and predicts bounding boxes within each grid. These bounding boxes are presented in the format (x, y, w, h, C, $p(c_1)$, $p(c_2)$, …), where (x, y) represents the center coordinates of an object, and (w, h) represents its width and height. C is the confidence coefficient for detection, and ($p(c_1)$, $p(c_2)$, …) show probabilities for different classes. A specific loss function evaluates differences between predicted and ground-truth bounding boxes. By optimizing neural network parameters through backpropagation to reduce the loss function value, the model gradually learns features from our dataset.

The loss function for our network is given as follows:

$$\text{Loss}_{\text{total}} = \alpha \cdot \text{Loss}_{\text{BOX}} + \beta \cdot \text{Loss}_{\text{OBJ}} + \gamma \cdot \text{Loss}_{\text{CLA}}$$

where α, β, and γ are weighting coefficients. LossBOX represents the location difference between prediction and ground truth. For YOLOv5, Complete Intersection over Union (CIOU) is used to obtain this distance.

$$\text{CIOU} = \text{IOU} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v$$

where Intersection over Union (IOU) is calculated as follows:

$$\text{IOU} = \frac{\text{Intersection}}{\text{Union}}$$

The Intersection for a given bounding box can be derived as:

$$\text{Intersection} = \min(x_2, x_2^{gt}) - \max(x_1, x_1^{gt}) \times \min(y_2, y_2^{gt}) - \max(y_1, y_1^{gt})$$

where the label "gt" means the parameter is given by ground truth, and symbols without "gt" are obtained by prediction. Then Union is obtained:

$$\text{Union} = (x_2 - x_1) \times (y_2 - y_1) + (x_2^{gt} - x_1^{gt}) \times (y_2^{gt} - y_1^{gt}) - \text{Intersection}$$

Here v, , and c are introduced to represent distance more accurately:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2$$

LossOBJ refers to the reliability of the predicted rectangle compared with the actual ground-truth box, and LossCLA denotes the difference between predicted and specific classes. The YOLOv5 model employs Binary Cross Entropy (BCE) loss to evaluate this variance:

$$\text{Loss}_{\text{BCE}} = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log p(i) + (1-y_i)\log(1-p(i))\right]$$

where n represents the number of pixels within the predicted rectangle. The binary label y indicates false or true with values of 0 or 1, respectively. The probability p(i) denotes the likelihood predicted by the model that a given pixel belongs to the specified region or class.

2.2. Network Optimization

Unlike multi-class detection tasks, our network is trained with only a single object class. Consequently, β is set to zero in the loss function for simplicity. On the other hand, considering that RSO images are generally small relative to the whole frame and their intensity distribution resembles a classic Gaussian function, we modify LossBOX using the Normalized Wasserstein Distance (NWD; Wang et al. 2022b). The new LossBOX is expressed as:

$$\text{Loss}_{\text{BOX}} = 1 - \text{NWD}(N_a, N_b)$$

where N and N_b represent Gaussian distributions of prediction and ground truth:

$$N_a = \mathcal{N}(\mu_a, \Sigma_a), \quad \mu_a = (x, y), \quad \Sigma_a = \text{diag}(w^2/4, h^2/4)$$

$$N_b = \mathcal{N}(\mu_b, \Sigma_b), \quad \mu_b = (x^{gt}, y^{gt}), \quad \Sigma_b = \text{diag}((w^{gt})^2/4, (h^{gt})^2/4)$$

$$\text{NWD}(N_a, N_b) = \exp\left(-\frac{W_2^2(N_a, N_b)}{C}\right)$$

The parameter C is determined from the dataset and set to 16 to optimally perform the training task.

For the neck structure, we employ an additional upsample layer to detect faint and small RSO images. Different from the fundamental P3, P4, and P5 layers, we substitute P5 with P2 in the detection header, resulting in a 40% reduction in model parameters while improving detection capability. The modified structure is illustrated in Figure 1 [Figure 1: see original paper].

3. Application

3.1. Observations

We utilize a dedicated RSO survey telescope for trial observations and analyze our network's efficiency based on raw data. Detailed telescope information is shown in Table 1 .

During observation, considering the dynamical characteristics of high-Earth orbital RSOs, their angular movement is slow relative to a ground observer. Hence, the stare mode is adopted, meaning the telescope maintains pointing at a specific horizon field during exposure with the drive turned off. The exposure time for a single frame is 2 s, and consecutive images are acquired at the same azimuth and elevation. With this strategy, background stars appear as short streaks in images while RSOs in high-Earth orbital regions appear as points. Due to the relatively wide field, several RSOs may appear in the same image, as shown in Figure 2 [Figure 2: see original paper].

The time interval for switching fields is 2 minutes, yielding approximately 15 raw CCD frames per field when considering CCD camera readout time and telescope pointing setup. Observations were performed over 4 nights. After acquiring large volumes of raw CCD images, object detection and tracklet extraction are performed using previous algorithms. Here, "tracklet" refers to a series of measurements for the same object, including centroids and observed positions in right ascension and declination. In detail, sources in each image are extracted and their measurement positions obtained, then equatorial positions are derived through astrometry. Based on movement differences between stars and RSOs in consecutive images, RSO candidates are detected and tracklets generated. Finally, these tracklets are correlated with the catalog (Yu et al. 2021). Tracklets not correlated with the catalog are recognized as false detections, while correlated tracklets serve as a priori information for model training and validation. After optimization, our previous reduction pipeline can detect RSOs in near real-time. It should be noted that after tracklets are extracted with our developed network, orbit correlation is also performed, yielding the numbers of false detections and correlated objects to evaluate network performance.

Detailed results from the previous method are shown in Table 2 , including observed fields, raw images acquired each day, extracted tracklets, and uncorrelated and correlated objects after reduction. The distribution of eccentricity and semimajor axis for correlated RSOs is shown in Figure 3 [Figure 3: see original paper]. Approximately 9000 raw images can be obtained each day, with more than 2800 tracklets extracted. It should be noted that one tracklet generally includes 8–15 position measurements, making the extensive a priori detections feasible and reliable for CNN application.

3.2. Dataset Construction

The obtained dataset provides RSO information including position measurements and image sizes. Artificial intelligence models can learn features from this dataset and perform object detection on new data, making high-quality dataset construction crucial. Generally, measured positions (x, y) and dimensions (width w, height h) require manual annotation. In our test based on previous object detection methods, we have already obtained RSO measured positions. Another issue is choosing optimal box sizes for the dataset. RSO image widths and heights are obtained with SExtractor using a 3σ threshold, with distributions shown in Figure 4 [Figure 4: see original paper]. The majority of RSO images (approximately 85%) are smaller than 20×20 pixels, differing from background stars which exhibit high pixel values and larger sizes. Considering the whole image size is 2048×2048, these are typically small and faint objects. Based on this distribution and after investigation, both width and height are set to 16 pixels to achieve a balance between performance and training efficiency.

Variations in background levels and object brightness can cause confusion within the model, making data normalization crucial. To address this challenge, we apply a gray transformation to augment faint RSO image signals:

$$I_{\text{new}} = \frac{I_{\text{orig}} - \text{bkg}}{\sigma}$$

where bkg and σ are the mean background level and standard deviation obtained by SExtractor, respectively. The augmentation effect is shown in Figure 5 [Figure 5: see original paper].

3.3. Model Training

Although model parameters have been optimized, the YOLO network still contains more than six million parameters. The training process involves adjusting these parameters to task-specific characteristics. Starting from zero or random values would require many more epochs for parameter convergence, consuming substantial time on large datasets. Parameter initialization can accelerate this process and reduce training epochs, so instead of training on the entire dataset initially, we create a smaller dataset for pre-training to obtain initial weights for our RSO detection task. This dataset consists of 200 randomly selected images based on RSO images from the first day, with a 3:1 training-to-validation ratio. All these images are manually labeled to avoid abnormal conditions like streak presence or dense star fields. Pre-training requires approximately 1 hour, yielding initial weights for subsequent model training. With pre-trained weights, a model trained on one day of data converges in approximately 50 epochs, requiring about 10 hours on a personal desktop with an NVIDIA RTX 3060 (6G), compared to 40 hours without pre-training. The time cost would reduce to approximately 2 hours using a professional workstation with multiple GPUs. The loss function for initial weight training is shown in Figure 6 [Figure 6: see original paper], along with precision and recall curves, indicating that the loss function decreases significantly and converges after 150 epochs while precision and recall rates approach 1, achieving warm-up for subsequent training.

To analyze network performance, we use two different training and testing strategies for comparison. In the first strategy, we train the model on day 1 data and test on day 2 data, then train on day 3 data and test on day 4 data. Each dataset includes more than 7000 images and 35,000 labels with a 3:1 training-to-validation ratio. In the second strategy, data from the first two days are used for training and data from the remaining two days for testing. The convergence epoch remains at 40 for both strategies, but training time increases for the second strategy due to doubled data size. Figure 7 [Figure 7: see original paper] demonstrates that pre-trained weights work effectively, with the training process starting from low loss and high recall, saving time costs.

4. Results and Discussions

After network training, we evaluate performance on the test set and analyze new objects detected only by our network to investigate improvements. In the object detection phase, after obtaining detection results for each frame, we use a breadth-first search strategy to generate measurement positions for all potential candidates with their confidence levels. To minimize false alarms, we select the first eight candidates based on confidence ratio. Coupled with the well-trained model, this approach promotes detection efficiency.

With the two strategies mentioned previously, consecutive detection results are extracted for each sky field, tracklets obtained, and orbit correlation performed. Results are shown in Tables 3 and 4 , respectively. Both training strategies yield more correlated tracklets and detected objects than the traditional method, demonstrating superior performance.

Furthermore, with pre-trained weights, our network completes object detection in less than 0.1 s for a 2048×2048 frame, not accounting for file I/O time. In comparison, the traditional method requires 1–2 s for the detection phase. This implementation benefits from CNN structure: if frame size increases dramatically as is likely in the future, time cost will not extend significantly, whereas traditional image processing methods would experience multiplied time costs.

Further investigations use data from the second strategy, which extracts more RSOs and tracklets with fewer false detections. The magnitude distribution of detections from the second day, obtained with both traditional methods and our network, is shown in Figure 8 [Figure 8: see original paper]. The distributions are similar, peaking between 10–12 mag with the faintest detections exceeding 16 mag. Our method achieves slightly better performance for both bright and faint objects. Details are shown in Table 5 . New RSOs are detected in images from all four days. Some tracklets and objects are detected exclusively by either our network or the previous traditional method, making the detections complementary. Images of newly detected objects are shown in Figure 9 [Figure 9: see original paper], indicating these detections appear not only as point-like sources but also as streaks or irregular shapes. After calculating their apparent angular velocity, they are confirmed as typical RSOs in high-Earth orbital regions. The semimajor axis and eccentricity of these objects, obtained through orbit determination, are shown in Figure 10 [Figure 10: see original paper], demonstrating that most newly detected objects are in highly elliptical orbits (HEOs) with unique dynamical characteristics. With data fusion, the number of tracklets and objects improves without deploying new infrastructure, enhancing system efficiency.

5. Conclusions

Due to modern optical and electronic technology development, RSO survey infrastructure and instruments have greatly improved, creating challenges for traditional data reduction methods when handling large data volumes. Here we develop a novel object detection technique using deep learning. Based on an end-to-end deep learning framework, our model is both trained and tested on large volumes of raw data. Results demonstrate performance comparable to previous methods with real-time data reduction capability. Additionally, our network detects numerous new tracklets and objects, improving system efficiency. This work shows that AI-based techniques can achieve strong synergy with astronomical data reduction and will certainly play vital roles in the future.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (NSFC, grant Nos. 12473079 and 12073082), and the National Key R&D Program of China (No. 2023YFF0725300). The authors thank the anonymous reviewers and Editor-in-Chief for valuable comments that substantially improved the manuscript.

References

Bertin, E., & Arnouts, S. 1996, ApJS, 117, 393
Du, J., Hu, S., Chen, X., & Cao, H. 2022, MNRAS, 511, 3377
Dumitrescu, F., Ceachi, B., Truică, C.-O., et al. 2022, Aeros, 9, 520
Grishin, K., Mei, S., & Ilic, S. 2023, A&A, 677, A101
He, Y., Wu, J., Wang, W., Jiang, B., & Zhang, Y. 2023, PASJ, 75, 1311
He, Z., Qiu, B., Luo, A. L., et al. 2021, MNRAS, 508, 2039
Hickson, P. 2018, AdSpR, 62, 3078
Jia, P., Liu, Q., & Sun, Y. 2020, AJ, 159, 212
Kouprianov, V. 2008, AdSpR, 41, 1029
Matney, M. J., Stansbery, E., Africano, J., et al. 2004, AdSpR, 34, 1137
Molotov, I., Agapov, V., Titenko, V., et al. 2008, AdSpR, 41, 1022
Schechter, P. L., Mateo, M., & Saha, A. 1993, PASP, 105, 1342
Schildknecht, T. 2007, A&ARv, 14, 41

Stetson, P. B. 1987, PASP, 99, 191
Sun, R., & Jia, P. 2017, PASP, 129, 044502
Sun, R. Y., Yu, S. X., Zhao, C. Y., et al. 2019, PASJ, 71, 67
Sun, R. Y., Zhan, J. W., Zhao, C. Y., et al. 2015, AcAau, 110, 9
Sun, R. Y., & Zhao, C. Y. 2013, RAA, 13, 604
Sun, R. Y., Zhao, C. Y., & Lu, Y. 2016, RAA, 16, 29
Tingay, S. J., Kaplan, D. L., McKinley, B., et al. 2013, AJ, 146, 103
Wang, F., Ge, J., & Willis, K. 2022a, MNRAS, 516, 5785
Wang, J., Xu, C., Yang, W., & Yu, L. 2022b, arXiv:2110.13389
Yanagisawa, T., Nakajima, A., Kadota, K.-I., et al. 2005, PASJ, 57, 399
Yu, S. X., Wang, X., & Zhu, T. L. 2021, AdSpR, 68, 71
Zhang, B., Hu, S., Du, J., et al. 2024, PASP, 136, 054502
Zhang, C., & Zhao, C. Y. 2021, AnABC, 93, e20200827
Zhang, Z. P., Yang, F. M., Zhang, H. F., et al. 2012, RAA, 12, 212

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.