Check for updates

Intelligent Intelligence Perception Technology and Applications Based on space-air-ground Multi-Modal Data Fusion

Zhuo Lin[1,2]

[1] Fujian Institute of Scientific and Technological Information, Fuzhou, Fujian,China
[2] Fujian Key Laboratory of Information and Network, Fuzhou, Fujian,China
linchamp@163.com

Abstract. With the continuous advancement of information technology, aerospace technology has become one of the crucial directions in informatization construction. Intelligent intelligence perception based on the fusion of space-air-ground multi-modal data refers to the acquisition of data through means such as satellites, drones, ground sensors, and open source data, followed by effective fusion and analysis using intelligent technologies. This enables automatic detection, tracking, identification, and early warning of specific targets and tasks, providing decision-makers with valuable comprehensive intelligence information. This study constructs a full-process application service model that includes "unstructured data sharing platform + multi-modal data fusion + big data intelligence analysis + situation monitoring and early warning system". Then, the study integrates space-air-ground data fusion with the power inspection application field, takes the target detection and recognition of power facilities such as substations and power transmission tower as cases, and explores the intelligent perception of power facilities based on the multi-scale remote sensing image object detection technology of Faster R-CNN. Finally, it contemplates the application scenarios for intelligent intelligence perception based on space-air-ground data fusion with other core technologies, including the Space-air big data service center, the remote sensing monitoring platform for major scientific and technological industrial projects, the "one-map" intelligence monitoring platform for science and technology industrial projects, and the terminal intelligent intelligence perception service integrating large language models.

Keywords: Space-air-ground Big Data, Intelligent Intelligence Perception, Target Detection; Power Facilities

1 Background

With the continuous advancement of information technology, aerospace technology has become one of the crucial directions in informatization construction. It is predicted that the global market for satellite remote sensing services will experience rapid growth in the coming years and the potential is immense, with governments around the world

© The Author(s) 2024
C. Bai et al. (eds.), *Proceedings of 2023 China Science and Technology Information Resource Management and Service Annual Conference (COINFO 2023)*, Advances in Economics, Business and Management Research 293,
https://doi.org/10.2991/978-94-6463-498-3_3

increasingly investing in the application of space-air-ground big data[1]. High-resolution Earth observation systems are one of the important development directions of aerospace technology. These systems, utilizing satellites, aircraft, and stratospheric balloons as platforms, enable high-resolution observations of the Earth's surface, providing more detailed and accurate information for environmental resource surveys, urban planning, agricultural production, and other fields[2]. Moreover, with the rapid progress of remote sensing and computer technology, obtaining high-resolution remote sensing images is no longer a challenge. However, the current pressing issue in both academic and industrial circles is how to quickly, automatically, and intelligently extract target information from the collected remote sensing images[3].

Therefore, the intelligent intelligence perception proposed in this study, based on the fusion of space-air-ground multi-modal data, refers to the acquisition of data through means such as satellites, drones, ground sensors, and open-source data (primarily focusing on science and technology data, public sentiment data, and security data). These data are characterized by multi-modal, heterogeneous, complex, and distributed nature, necessitating the use of intelligent technologies for effective integration and analysis. This enables the automatic detection, tracking, identification, and early warning of specific targets and tasks. Such an approach can be applied in various fields, including national defense, science and technology, national security, natural resources, and urban planning, providing decision-makers with comprehensive and valuable intelligence information[4].

2 Overall Framework and Core Technology

2.1 Overall Framework

This study constructs a comprehensive application service model covering the entire process of "unstructured data sharing platform + multi-modal data fusion + big data intelligence analysis + situational monitoring and early warning system." The unstructured data sharing platform provides a secure and convenient way to access and utilize space-air big data; the multi-modal data fusion is aimed at multiple heterogeneous data sources to achieve thematic aggregation of multi-modal information; the big data intelligence analysis primarily uses intelligent technologies to extract useful information and intelligence from massive data, providing real-time and accurate data support; the decision-making situational monitoring and early warning system mainly achieve intelligent tracking and early warning for specific tasks through continuous situational monitoring, providing decision-making support.

2.2 Core Technology

Multi-scale object detection in remote sensing images based on Faster R-CNN. Utilizing a feature extraction network for image feature extraction, high-level feature maps are sampled through a multi-level fusion structure, which are then fused with lower-level feature maps to generate a series of different-scale feature maps. Region

Proposal Network (RPN) is used on these different-scale feature maps to generate candidate regions. Finally, a classification and regression network maps candidate regions of different scales to the corresponding feature maps for classification and position regression, thus yielding the ultimate object detection results[5].

The remote sensing image object recognition technology based on YOLO-V7. Enriching the semantic information module on top of YOLO-V7, by analyzing and processing contextual information surrounding the targets, enhances recognition accuracy and robustness. Additionally, incorporating activation functions more suitable for object detection tasks to capture target features elevates detection precision. By refining the detection end structure within the YOLO-V7 network architecture, it achieves high-precision intelligent identification of various small targets. Coupled with fine-tuning parameters using a large language model, in tasks such as object detection, tracking, and recognition, employing end-to-end reasoning allows the model to extract and focus on images most relevant to instructions, thereby enhancing its intelligence and perceptual capabilities for actionable insights.

Integrating open-source intelligence perceptual technology powered by large language models. Building upon conventional intelligent technologies, conducting in-depth correlation analysis of massive multi-modal open-source intelligence, extracting and structuring entities, relationships, and attributes, to develop an open-source intelligence perceptual engine centered around knowledge computation. Integrating large language models into intelligence perception enables rapid extraction of vital information from data sources, fine-grained knowledge extraction, and automated generation of domain-specific knowledge graphs. This facilitates the intuitive presentation of complex associations and allows for the exploration of intelligence through interactive question-answering methods, supporting applications such as analysis, monitoring, and early warning systems. The system delivers efficient and high-value intelligence support and knowledge services.

3 Research Case

To further elucidate the application, this study combines satellite-terrestrial data fusion with the field of power inspection, using substations and power transmission tower facilities as case studies for target detection and recognition. Based on Faster R-CNN, we employ remote sensing image multi-scale target detection techniques to explore intelligent analysis models for power facility data intelligence. This integration approach is beneficial for applying remote sensing imaging technology to real-world scenarios, meeting the needs of the power industry for intelligent remote sensing image analysis technology. Furthermore, by integrating the knowledge base of power facilities, we aim to enhance the safety level and management efficiency of power facility operations.

3.1 Platform Design

To build a high-quality web product, we utilize the Django framework. For the front end, we employ the CSS and JavaScript framework Bootstrap. The back end runs on the Web server Nginx, and we utilize both SQLite3 and MySQL as databases. Both databases fulfill the functionalities required by the website, and through Django's ORM framework, we manage the data, meeting users' demands for data analysis and providing efficient, reliable data analysis services.

On the front end, users can log in to the website through a browser to perform data analysis operations. The website is presented using a combination of HTML, JavaScript, and CSS, and the Bootstrap framework is employed to achieve a modern web interface.

In the back end, when a user sends a static HTTP request, the request is handled by the backend Nginx web server. When a user sends a dynamic request, Nginx forwards the request to the underlying uWSGI for processing. uWSGI acts as an intermediary between Nginx and Django, parsing HTTP requests and passing them on to Django. Django receives WSGI requests and returns HTTP responses, which uWSGI then forwards to Nginx, which in turn sends them to the user's client[6].

In terms of databases, Django uses its ORM framework to manage data. The data structure is described using models, facilitating easy data manipulation and management.

3.2 Multi-scale Object Detection in Remote Sensing Images

Faster R-CNN primarily consists of three components: the feature extraction network, the Region Proposal Network (RPN), and the classification and regression network. In convolutional neural networks, if there are many consecutive convolutional layers that output feature maps of the same size, these layers are considered to be at the same network level. During the detection process, the feature extraction network is first employed to extract features from the image. This is achieved by upsampling feature maps obtained from higher levels and fusing them with feature maps from lower levels, resulting in a series of feature maps at different scales (F5、F4、F3 and F2). Subsequently, the RPN is used to generate candidate regions on these feature maps at different scales. Finally, the classification and regression network maps the candidate regions at different scales to their corresponding feature maps for classification and localization regression, yielding the final target detection results[5].

3.3 Focused Grid Distribution for Rapid Target Location Search

To extract features from the dataset, we utilize neural networks such as ZF, VGG-16, and ResNet-50. Due to the varying scales of remote sensing image targets, it is necessary to set appropriate candidate regions on the last feature maps of the aforementioned networks when generating candidate regions with RPN[8]. The corresponding detection accuracies (%) are shown in Table 1. From Table 1, it is evident that the choice of the feature extraction network and the setting of the candidate region scales significantly impact the detection accuracy of remote sensing image targets. When the candidate region scale matches the original Faster R-CNN candidate region scales (128×128, 256×256, 512×512), all three feature extraction networks achieve the worst detection accuracy. This is because the original candidate region scales are designed for natural image targets, which are generally larger than remote sensing image targets, making them unsuitable for remote sensing image target detection. In Table 1, both the ZF and VGG-16 networks achieve their best detection accuracy at candidate region scales of (64×64, 128×128, 256×256, 512×512), with values of 78.39% and 80.55%, respectively. For ResNet-50, the best detection accuracy of 88.89% is achieved at candidate region scales of (32×32, 64×64, 128×128, 256×256). Even when smaller candidate regions are set for ZF and VGG-16, their weaker network feature extraction capabilities compared to ResNet-50 result in more limited features being extracted from smaller candidate regions, leading to lower accuracy in extracting small targets. Therefore, when setting RPN candidate region parameters, we use the scales of the candidate regions that achieved the best detection accuracy for ZF, VGG-16, and ResNet-50, respectively. For the three improved networks, specific parameter settings are detailed in Table 2.

Table 1. Comparison of detection accuracy of candidate regions with different scales

Candidate Region Scale	Feature extraction etwork		
	ZF	VGG-16	ResNet-50
(32×32, 64×64, 128×128)	78.17%	80.08%	88.87%
(64×64, 128×128, 256×256)	78.32%	80.44%	85.66%
(128×128, 256×256, 512×512)	70.52%	79.38%	80.76%
(32×32, 64×64, 128×128, 256×256)	78.09%	80.32%	88.89%
(64×64, 128×128, 256×256, 512×512)	78.39%	80.55%	86.81%
(32×32, 64×64, 128×128, 256×256, 512×512)	78.32%	80.45%	88.76%
(32×32, 64×64, 128×128)	78.17%	80.08%	88.87%

Table 2. Improved Network Candidate Region Scale Settings

Feature Map	Resolution Level	Candidate Region Scale
F2	High resolution	32×32
F3	Higher resolution	64×64
F4	Lower resolution	128×128
F5	Low resolution	256×256
F2	High resolution	32×32

3.4 Detection Result

To evaluate the effectiveness of the algorithm proposed in the research for target detection in remote sensing images, two widely used standard metrics methods: the Precision-Recall Curve (PRC) and the Average Precision (AP) are employed as the evaluation criteria for target detection. The PRC plots precision against recall, which represents the relationship between precision and recall values as the threshold changes. The AP is the average of precision values as recall varies from 0 to 1, which is essentially the area enclosed by the PRC curve and the axes. The specific calculation formulas for precision and recall are as follows:

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

In the formula, TP (True Positive) represents a predicted target that is actually a target, FP (False Positive) represents a predicted target that is actually background, and FN (False Negative) represents a predicted background that is actually a target. In this paper, prediction boxes with an IoU (Intersection over Union) value greater than or equal to 0.5 with the target label box are considered TP, otherwise, the prediction box is classified as FP.

The evaluation of the detection model accuracy for power facility targets in unstructured image data of power transmission poles and specific regional targets of substation-like power facilities was conducted (see **错误!未找到引用源。** and **错误!未找到引用源。**). The results indicate that the accuracy of the model for power transmission pole-like unstructured image data power facility targets (mAP50) is 0.886, while the accuracy of the model for substation-like power facility specific regional targets (mAP50) is 0.821. These results suggest that our target detection model can achieve high accuracy in different scenarios and is capable of effectively detecting power facility targets and specific regional targets. These findings are of significant importance for the monitoring and maintenance of power facilities, as they can enhance the safety and stability of power infrastructure.

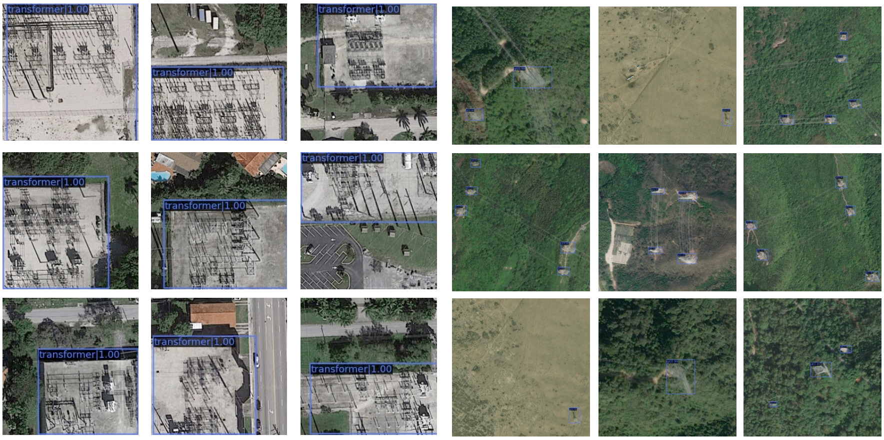

Fig. 1. Models detection results

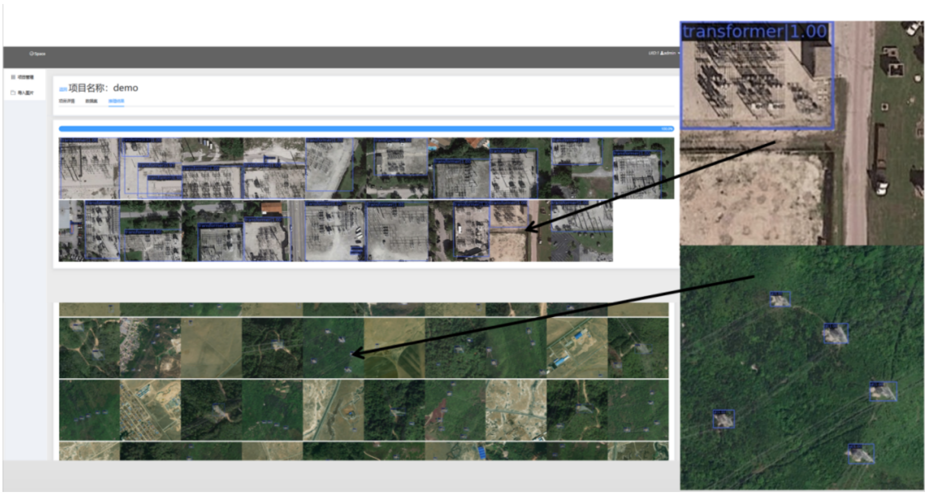

Fig. 2. Overall effect of platform detection results

4 Discussion and Conclusion

4.1 Intelligence perception application for power inspection scenarios based on space-air-ground data

Utilizing intelligent recognition and intelligence perception from space-air-ground data can bring multiple benefits to the power industry. Firstly, this research can quickly identify issues in the distribution of power grids, such as equipment failures and damage, allowing for timely maintenance and handling, which significantly improves the efficiency of power grid maintenance, reduces maintenance time and costs, and enhances the operational efficiency of power grid facilities. Secondly, it is possible to obtain panoramic images of power grid targets with high resolution, making power grid data more visualizable and analyzable, which increases the safety and stability of power grid facilities, thereby enabling better management and decision-making for power grid assets. Lastly, conducting related research can drive the application of integrated intelligent technology for space-air-ground data fusion in the power industry, offering more efficient, accurate, and reliable remote sensing image intelligent analysis services, and thus propelling the digital transformation and intelligent development of the power industry.

4.2 Intelligent intelligence perception application scenarios of space-air-ground multi-modal data fusion

To cater to various need, intelligence perception technologies based on space-air-ground multi-modal data fusion can cater to a multitude of domains, including national security, science and technology security, urban security, environmental monitoring, among others. With a certain level of innovativeness, the key applications scenarios primarily encompass:

The Space-air Big Data Service Center. By leveraging the big data from space-air platforms, a space-air data service center is constructed to provide a vast and stable foundation of space-air big data for the development of various fields. Specific product and service formats include satellite regional mapping, global mapping services, support for custom-selected area data downloads, compatibility with other image services following OGC (Open Geospatial Consortium) standards, and the provision of multi-period data axis views for specific locations, multi-period image roll comparison, and transparency overlay viewing, among other features.

The Remote Sensing Monitoring Platform for Major Science and Technology Industrial Projects. By utilizing satellite remote sensing combined with intelligent interpretation techniques, this platform dynamically monitors major industrial projects. It analyzes the land use status or phenomenon changes in the same area through multi-temporal and multi-spectral image recognition. The platform assesses the dynamic changes and future development trends in terms of spatial distribution and quantity, enabling the timely and accurate grasp of the construction status of major science and technology industrial projects. It constructs a diverse service model for remote monitoring of industrial projects and provides decision-making support for the management of science and technology industrial projects.

The "One-map" Intelligence Monitoring Platform for Science and Technology Industrial Projects. By integrating satellite remote sensing data, low-altitude aircraft data, Internet of Things (IoT) data, and open-source data, an "All-in-One" map-based intelligence monitoring platform for industrial projects is formed. Centered around data collection, big data processing, and artificial intelligence decision support, this platform provides real-time, panoramic space-air big data combined with multi-source, multi-modal data intelligence from vertical sectors. It effectively manages and controls the progress of major science and technology industrial projects at all stages, offers intelligence monitoring and early warning, and provides decision-support for leaders and managers.

The Terminal Intelligent Intelligence Perception Service with Large Language Models. This service combines satellite remote sensing technology with open-source intelligence, utilizing large-scale image recognition of space-air data to elevate intelligence monitoring to a space-air perspective. It traced the origins of open-source intelligence and evaluates the authenticity and reliability of significant intelligence. By leveraging large language models, the service enhances the entire process from data collection, entity extraction, intelligent analysis, attitude and position analysis, to thematic reports, providing services such as public sentiment perception, situation warning, event tracking, network structure evolution analysis, and intelligent recommendation of behavioral strategies. This realization achieves the cross-border integration of satellite remote sensing technology and open-source intelligence.

5 Acknowledgement

The paper is supported by the Basic Research Project of Public Welfare Scientific Research Institutions in Fujian Province in 2021, "Research on Technology Innovation Service Application Technology Based on Big Data Intelligent Analysis and Mining" (Project Number: 2021R1008006), the Basic Research Project of Public Welfare Scientific Research Institutions in Fujian Province in 2022, "Construction and Application of Taiwan District Technology Innovation Information Monitoring Platform" (Project Number: 2022R1008001), and the Innovation Fund Project in Fujian Province in 2023, "Research and Application of an Intelligent Grid Loss Reduction and Carbon Emission Reduction Method Based on Multi-algorithm Cross-validation and Multi-model Integration" (Project Number: 2023C0026). It is one of the research results of these projects.

References

1. Deren,Li: Towards Geo-spatial Information Science in Big Data Era. Acta Geodaetica et Cartographica Sinica 45(04), 379-384 (2016).
2. Deren Li, Mi Wang, Xin Shen, Zhipeng Dong: From Earth Observation Satellite to Earth Observation Brain. Geomatics and Information Science of Wuhan University 42(02), 143-149 (2017).
3. Mingsong Yuan: Target Recognition and Location Change Analysis System based on Remote Sensing Image. Harbin Institute of Technology University (2021).
4. Wu Xue, Ling Zhao, Peng Wang: Satellite-earth Coordinated On-orbit Intelligent Target Recognition of Optical Remote Sensing Images. Spacecraft Recovery & Remote Sensing 42 (03), 128-137(2021).

5. Miaomiao Sha, Yu Li, An Li: Multiscale aircraft detection in optical remote sensing imagery based on advanced Faster R-CNN. National Remote Sensing Bulletin 26 (08), 1624-1635(2022).
6. Yue Zhou: Django-Based Campus Print Reservation Website Design and Implementation. Jilin University (2013).
7. Zhiyou Wang, Huan Li, Zizeng Liu, Jiamin Wu, Zuxian Shi: Satellite Image Change Monitoring Based on Deep Learning Algorithm. Computer Systems & Applications 29 (01), 40-48(2020).
8. Xiujuan Yu, Xufeng Pang: Dynamic Monitoring of Land Use Based on Multi-source High-resolution Remote Sensing Images. Technology Innovation and Application 13 (02), 16-19(2023).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.