Journal of Innovation and Development
ISSN: 2958-4752 | Vol. 9, No. 1, 2024

A Method for Generating Intelligent Message Products Based on Intelligent Interpretation

Yang Liu

The 27th Research Institute of China Electronics Technology Group Corporation, Zhengzhou, Henan, 450047, China

Abstract: This article proposes a method for generating intelligent message products based on data intelligent interpretation, including pre-configured module, data analysis modules, message modules, and formatted modules. Firstly, utilizing remote sensing data intelligent interpretation algorithms to extract scene elements. Secondly, complete message statement generation by generating decision trees. Finally, example applications have shown that proposed method can solve the problems of complex processes, high labor costs, and low system timeliness in the production process of standard message products, achieve product production automation, and improve data application efficiency.

Keywords: Intelligent interpretation, Knowledge graph, Decision tree, Intelligent message.

1. Introduction

With the rapid development of satellite constellations and satellite Internet constellations, the number of satellites will rise sharply. The continuous enrichment of satellite data sources and the continuous expansion of new load spectrum make it possible to obtain tens of millions of square kilometers of data every day [1]. In addition, remote sensing satellite technology is gradually maturing, promoting the creation of various types of high-quality remote sensing satellite data products in military, civilian, commercial and other fields, gradually creating considerable application value. The data interpretation and analysis industry relying on a large number of professional personnel has also gradually formed, divided into two types: manual visual interpretation and semi-automatic interpretation, which mainly rely on accumulated expert knowledge for interpretation. Liu Jun et al. introduced the image features of major vegetation types and methods for vegetation interpretation in remote sensing images [3]. Shu Jianwen et al. proposed a road detection method in aerial remote sensing images based on principal component analysis algorithm, and used LBP method for feature extraction to improve the accuracy of detection [4].

Although the existing manual interpretation of remote sensing data has the characteristics of high accuracy and strong reliability, there are problems such as poor timeliness and low data utilization. With the geometric growth of remote sensing product data volume, traditional data interpretation and analysis methods are no longer able to meet the real-time interpretation and analysis needs of massive data. Especially in the application of remote sensing product data [5]， its application process is very similar in industries such as military, environmental protection, meteorology, geology, and public management. Users need professional personnel to first interpret and analyze remote sensing product data to extract key element attributes of interest. After analyzing the elements to form results, according to fixed format requirements, the analysis results are then converted into official documents or messages and submitted to decision-making agencies for assistance in making decisions[6].

However. Currently, relying on manual labor to complete related tasks has problems such as low work efficiency and high labor costs, which can no longer better adapt to the high timeliness and application requirements of big data. To solve the above problems, data intelligent interpretation technology has emerged and has become the main research direction of remote sensing profession [7] [8]. Among them, artificial intelligence methods represented by deep learning can significantly improve the efficiency of remote sensing data interpretation. They mainly use deep neural networks to perform end-to-end feature extraction on input remote sensing data, in order to accurately and quickly achieve various interpretation tasks such as land cover classification, change detection, and 3D reconstruction. Hu et al. first applied Convolutional Neural Network (CNN) to the problem of hyperspectral remote sensing image classification [9]. Chen et al. used multi-layer convolutional neural networks and utilized the proposed virtual sample enhancement method to improve classification performance and reduce dependence on a large number of training samples [10]. Liang et al. processed the deep features extracted by CNN by using a classification framework with sparse representation function, which can more intuitively demonstrate CNN's abstract representation of features [11].

Although data intelligent interpretation algorithms have gradually become a powerful tool in the hands of professional analysts, there are still problems in practical application scenarios such as remote sensing, data engineering, and public management where the threshold for multi-disciplinary integration is too high, which is not conducive to the efficiency of data application and industry development under multi-disciplinary integration. Therefore, this article proposes an intelligent message product generation method based on data intelligent interpretation. On the basis of intelligent interpretation and analysis algorithms, it uses scene template configuration, knowledge graph inference, and decision tree mechanism to provide standard message products that are closer to expectations, solving the problems of long interpretation time and high labor costs for massive data, and thus improving the efficiency of data application under multi professional integration.

2. Intelligent Message Product Generation Framework

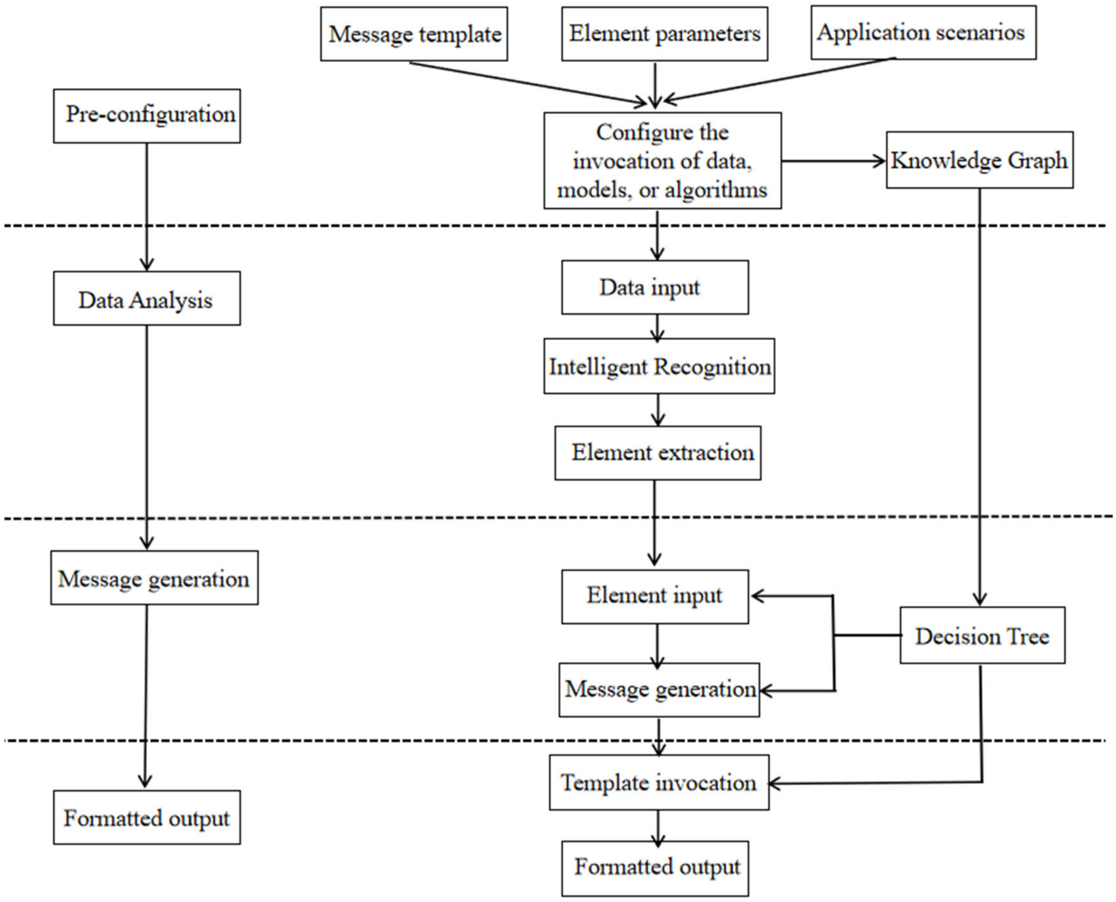

Figure 1. Overall flowchart of intelligent message product generation method based on data intelligent interpretation

The overall process of the intelligent message product generation method based on data intelligent interpretation mainly consists of four modules: pre configuration, data analysis, statement generation, and message generation. The pre configuration module pre configures application scenarios, element parameters, and message templates to construct a knowledge graph; The data analysis module completes data analysis through steps such as data input, intelligent recognition, and feature extraction, and draws relevant analysis conclusions; The sentence generation module uses intelligent interpretation and analysis of data to obtain sentence patterns and generate message sentences using knowledge graphs; The message generation module uses a knowledge graph to generate decision trees and complete the call of message output templates to achieve formatted output.

3. Method for Generating Intelligent Message Products

The intelligent message product generation method is based on remote sensing data intelligent interpretation algorithm, extracts scene elements, and uses decision tree generation to complete message statement generation, greatly reducing the overall production process time of professional fusion analysis message products, realizing product production automation, meeting the timeliness requirements of future multi scene message products, and enhancing the application value of remote sensing images. It mainly includes four modules: pre configuration, data analysis, statement generation, and message generation.

3.1. Pre configuration

Pre configuration mainly involves pre configuring application scenarios, element parameters, and message templates to construct a knowledge graph, as follows:

(1) In the pre configuration process of application scenarios, it mainly includes building scenario information, scenario association configuration, scenario data element conversion, and scenario automation services.

Constructing scene information includes: scene type, scene elements, available data sources for the scene, scene spatiotemporal scope, and business relationships;

Scene information association configuration, configuring the association relationship between scene information and message templates;

Scene data element transformation refers to the data element transformation of scene information and its associations into "entities, relationships, and attributes" and importing them into a knowledge graph;

Scenario automation service, select the running scenario, start the scenario automation service, can call the configured scenario and automate its operation.

(2) Pre configure element parameters, including constructing element configurations and converting element data elements.

Constructing element configuration includes element types, element attributes, element relationships, and other related content;

Element data element transformation involves transforming elements into "entities, relationships, and attributes" and importing them into a knowledge graph.

(3) The pre configuration process for message templates includes parsing message templates, constructing message template configurations, semantic segmentation and annotation of message templates, and conversion of message data elements.

Analyzing message templates involves importing multiple application scenario type message templates, semantically segmenting and annotating the message templates to obtain message statements, sentence weights, sentence elements, and message formats;

Message association configuration, configuring the association relationship between message statements, statement weights, statement elements, and message formats for message scenarios;

The conversion of message data elements involves transforming the message into "entity, relationship, and attribute" data elements and importing them into a knowledge graph.

3.2. Data analysis

Data analysis is mainly completed through steps such as data input, intelligent recognition, and feature extraction to draw relevant analysis conclusions.

(1) In the data input stage, a list of data resources supporting analysis can be obtained from the data management system according to the scenario configuration, and data download caching and preprocessing can be carried out. The specific content is as follows:

Using satellite data to obtain the current month's data and the previous month's data list;

Conduct preliminary quality assessment of data to determine whether it meets the analysis requirements;

Download data to the data analysis service cache;

Perform image data pyramid construction processing, region slicing processing, denoising, and normalization data preprocessing procedures.

(2) In the stage of data intelligent recognition, target intelligent recognition algorithms can be called based on element configuration.

Establish a software service cluster, form datasets for different types of satellite payloads, train recognition models for specific target groups, and form a software service cluster;

Based on the scene attributes and associated element requirements, call the intelligent recognition service cluster and use data interpretation and recognition algorithms that are suitable for the scene satellite data type and element type to perform target recognition.

Push the target recognition results to the feature extraction service.

(3) After completing data intelligent recognition, it is necessary to extract elements and import them into the knowledge graph, as follows.

Obtain target recognition results and scene element features;

Extract feature attributes based on recognition results;

Convert elements into data elements of "entities, relationships, and attributes" and import them into a knowledge graph.

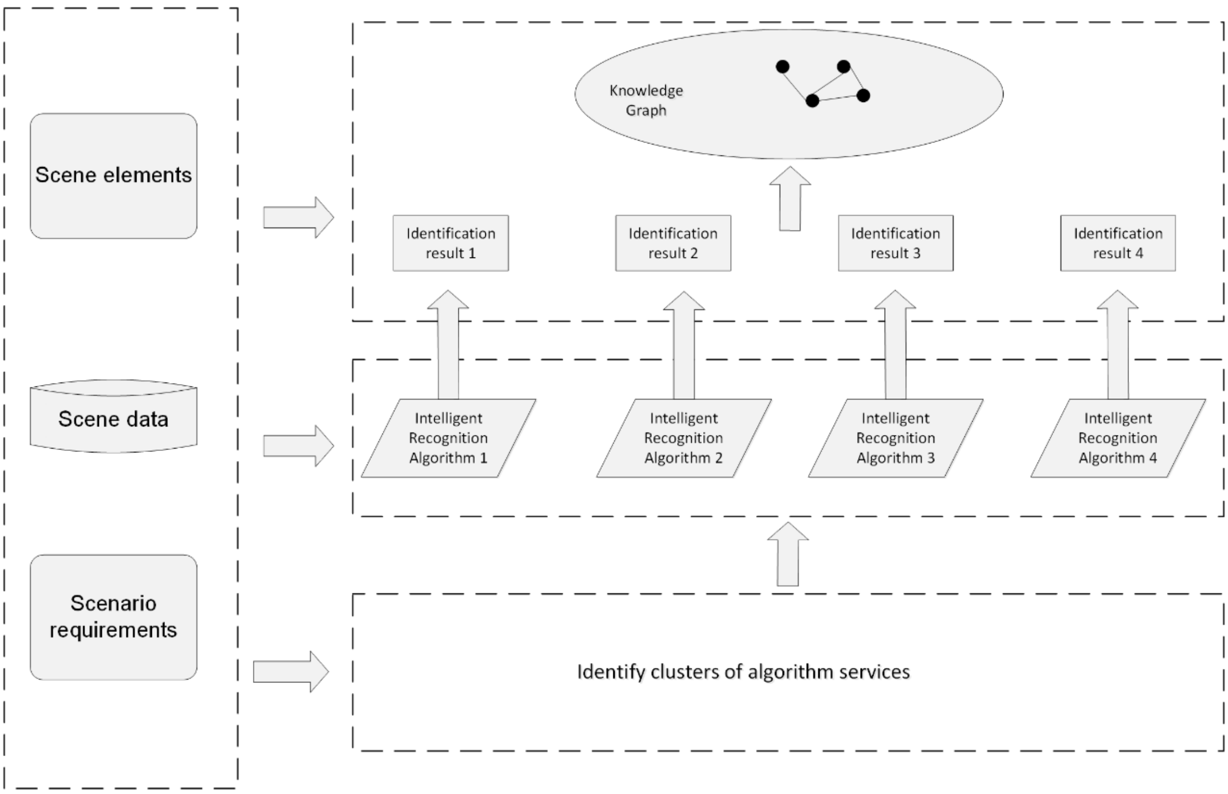

Figure 2. Working principle diagram of intelligent data interpretation

3.3. Statement Generation

Statement generation is mainly based on intelligent interpretation and analysis of data, using knowledge graphs to obtain sentence structures and generate message statements. It mainly includes two key steps: decision tree generation and message statement generation.

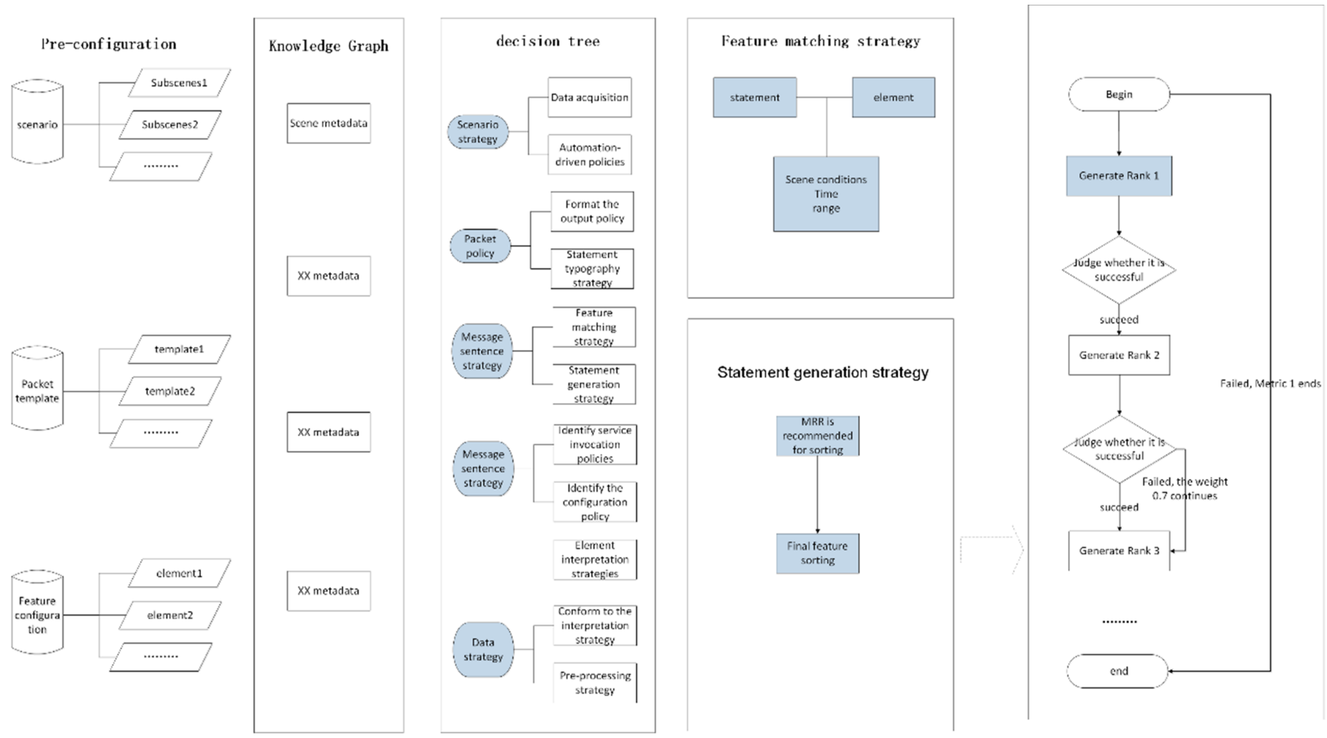

Figure 3. Decision tree generation process of intelligent message product generation method based on data intelligent interpretation

(1) When generating a decision tree, the knowledge graph is queried based on the scene configuration to obtain the Mean Recurrent Ranking (MRR) for matching elements. Then, the approximate elements are fused by combining the message statement weights and element fusion strategy, and the content with a message weight of 1 (required to be filled in) is sorted in advance to obtain the final element ranking, forming a message generation decision tree.

(2) Based on decision tree judgment, convert the results of element recognition and element analysis into message statement metadata.

3.4. Message Output

Message output is the use of knowledge graphs to generate decision trees, complete message output template calls, and achieve formatted output.

(1) Based on the metadata of message statements, call scene templates and replace statements with the same attributes and relationships in the template. When making template calls, the message product will be called based on the pre configured template and the relationship between statement elements.

(2) Firstly, based on actual needs, produce standardized formatted message products such as "briefings", "red headed official documents", etc; Then, based on the call message template, insert the generated message statements and charts into the corresponding content positions, and load the template font, paragraph, format, and other configurations; Finally, according to the message format requirements, call the production modules in formats such as doc, docx, pdf, rtf, etc.

4. Example Application

The intelligent message product generation method based on data intelligent interpretation proposed in this paper is applied to environmental governance scenarios using high-resolution XX four channel multispectral remote sensing satellite image data (5-meter resolution).

(1) Pre configure application scenarios, element parameters, and message templates.

Firstly, pre configure the application scenario. In the context of environmental governance, its refined application scenarios include water detection, greening detection, and bare soil detection. By configuring scenario relationships (water detection, greening detection, and bare soil detection belong to environmental governance scenarios), spatiotemporal scope (XX district of XX city, comparison of changes between this month and the previous month), loading strategy, and resource utilization strategy (e.g. analyzing once every 7 days and comparing data resources with high-resolution XX four channel multispectral remote sensing satellite image data from 30 to 60 days within 30 days), automated scenario driven capability is supported.

Secondly, perform pre configuration of element parameters. Taking the refined scenario of bare soil detection in environmental governance as an example, the main elements of configuration include bare soil judgment threshold, bare soil area, and governance measures. The parameters and weights of the elements are configured to form data elements of "entity, relationship, and attribute". The scene knowledge graph is imported to support subsequent algorithms to complete data analysis.

Finally, pre configure the message template. Configure message templates according to the actual usage needs of users. Generally, a complete list of message templates is entered during the system construction phase, and only selections are made during configuration, such as the "Monthly Report on Bare Soil Monitoring and Management in XX District, XX City". In addition, templates can be edited and modified through message editing tools to meet maintenance and expansion needs.

Specifically, in the configuration of message templates, weight values are assigned to message statements. When the weight value is 1, relevant content must be filled in, and when the weight value is less than 1, the conditions for filling in are determined from high to low; Associate the template content with the elements (for example, define the statement "The bare soil area in this area is XXX square meters" as the statistical bare soil area and associate it with the bare soil statistical elements), form data elements of "entity, relationship, attribute", and import them into the knowledge graph.

(2) Complete data input, intelligent recognition of remote sensing data, and feature extraction.

Firstly, complete the data input. According to the scenario configuration, use a knowledge graph to obtain and analyze the required data content, including spatiotemporal range, data type, satellite model range, etc. Obtain a list of data resources that support analysis from the data management system, and perform data download caching and preprocessing. For example, in the scenario of bare soil analysis, the steps are as follows:

1) According to the requirements of bare soil monitoring scenarios, use high-resolution XX satellite data to obtain a list of data for the current month and the previous month in XX district of XX city;

2) Conduct preliminary quality assessment of data (such as data volume, coverage range, cloud cover, etc.) to determine whether it meets the analysis requirements;

3) When the data meets the requirements, download the data to the data analysis service cache;

4) Perform data preprocessing procedures such as image data region slicing, pyramid construction, and denoising.

Secondly, achieve intelligent recognition of remote sensing data. According to the element configuration of bare soil monitoring scenarios, call the target intelligent recognition service, configure parameters such as confidence and recall based on the element configuration information, and input the preprocessed image slice data such as region slicing, pyramid construction, denoising, and normalization into the intelligent recognition service. For example, bare soil detection mainly relies on calling the high-resolution optical image satellite terrain recognition algorithm trained on YOLOv5, and using a 4 * 1024 * 1024 pixel window with 512 pixel steps to perform semantic segmentation and recognition of bare soil, governance measures, and other targets, in order to obtain recognized bare soil targets and governance measures targets. Due to the relatively regular nature of the land parcel data, different bare soil areas in the scene are spaced more than 10 meters apart. Therefore, the extreme coordinates of the target area can be used to form a rectangle and determine the overlap of the targets, and the target set can be fused and deduplicated.

Finally, implement feature extraction. The feature attributes associated with scenes in the knowledge graph are used to extract feature attributes from the recognition results. In the scenario of bare soil detection, the associated statistical analysis of bare soil elements involves measuring the area, dividing the area, matching measures, and conducting comprehensive statistics on the recognition results, forming a set of element data elements and an analysis result data element set, and importing them into the scene knowledge graph.

(3) Generate message statement metadata based on the decision tree

Firstly, query the knowledge graph based on the scenario configuration. The knowledge graph is embedded with a KGE measurement model, which combines element fusion strategies to form a message generation decision tree;

Secondly, based on decision tree judgment, the results of element recognition and element analysis are converted into message statement metadata. Specifically, the statement generation module includes two key steps: decision tree generation and message statement generation.

1) In the decision tree generation module, the knowledge graph is queried based on the scenario configuration to obtain Mean Recurrent Ranking (MRR) for the elements. Then, the approximate elements are fused by combining the message statement weights and element fusion strategy, and the content with a message weight of 1 (required to be filled in) is sorted in advance to obtain the final element ranking.

2) In the message statement generation module, the element information is interpreted into message statements based on the generated decision tree. For example, in the scenario of bare soil detection, during the production statistics monthly report, the data of each element (statistics: total bare soil area of 28k square meters...; plot 1: bare soil, area of 7k square meters, 01 district...) is converted into a set of natural statements such as "the existing total bare soil area of 28k square meters in XX district of XX city..." and "the coexistence of bare soil plots of 23.6k square meters in 01 district, including plot 1...".

(4) Generate message products by calling scene templates and outputting message formats.

1) Make template calls. According to the pre configured template, template calls are made to message products based on the relationship between sentence elements. For example, for the monthly report of bare soil inspection scenario statistics, if the message generation has completed the generation of "basic statistical information" (the total bare soil area in XX district of XX city is 28k square meters...), but has not generated "abnormal statistical information", the "bare soil monitoring monthly report" template is called to omit the abnormal statistical information statements with weights less than 1 in the monthly report, and complete the formatted message production.

2) Implement formatted output. Firstly, based on actual needs, produce standardized formatted message products such as "briefings", "red headed official documents", etc; Then, based on the call message template, insert the generated message statements and charts into the corresponding content positions, and load the template font, paragraph, format, and other configurations; Finally, according to the message format requirements, call the production modules in formats such as doc, docx, pdf, rtf, etc.

5. Conclusion

This article starts with the contradiction between the existing remote sensing image data interpretation and analysis methods and the real-time interpretation and analysis requirements of massive data, and proposes an intelligent message product generation method based on data intelligent interpretation. On the basis of intelligent interpretation and analysis algorithms, remote sensing data intelligent interpretation algorithms are used to extract scene elements. Simultaneously generate message statements by generating decision trees. Further examples demonstrate that the proposed method can solve the problems of complex processes, high labor costs, and poor timeliness in the production process of standard message products, and improve data application efficiency.

References

[1] Wan Wei, Zeng Yuan, Ren Xinyue, etc Reverse stitching chain strategy based on low orbit full coverage satellite constellation [J] Shanghai Aerospace (in Chinese and English),2023, 40(6):54-63.

[2] Wang Yi, Wu Xiaorui, Feng Huiying, etc Research and Practice of Model Driven Massive Data Interpretation System [J] Computer Measurement and Control, 2018, 26 (12): 5. DOI: CNKI: SUN: JZCK. 0.2018-12-043

[3] Liu Jun, Jia Xiuming Vegetation and soil interpretation in remote sensing images [J] Technology Intelligence Development and Economics, 2005, 15 (6): 3. DOI: 10.3969/j.issn.1005-6033.2005.06.080

[4] Shu Jianwen, Zhou Yunzhen Research and Simulation of Road Detection Methods in Aerial Remote Sensing Images [J] Computer Simulation, 2013, 30 (9): 4. DOI: 10.3969/j.issn.1006-9348.2013.09.049

[5] Cui Lu, Zhang Peng, enter the car Overview of Remote Sensing Image Classification Algorithms Based on Deep Neural Networks [J] Computer Science, 2018 (z1): 50-53

[6] Yao Yingmao, Jiang Xiaoyan Video pedestrian re recognition method based on graph convolutional network and self attention graph pooling [J] Computer Applications, 2023, 43 (3): 8. DOI: 10.11772/j.issn.1001-9081.2022010034

[7] Fu Kun, Wang Peijin, Feng Yingchao, etc Remote sensing cross modal intelligent interpretation: models, data, and applications [J] Chinese Science: Information Science, 2023, 53 (8): 1529-1559. DOI: 0.1360/SSI-2023-0055

[8] Yu Fei, Sui Zhengwei, Qiu Fengting, etc A review of the progress in constructing SAR image intelligent interpretation sample datasets [J] Network Security and Data Governance, 2023, 42 (S01): 97-105

[9] Hu,Wei,Huang,et al.Deep Convolutional Neural Networks for Hyperspectral Image Classification[J].Journal of Sensors, 2015.

[10] Chen Y , Jiang H , Li C ,et al.Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks[J].IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(10):6232-6251.DOI:10.1109/TGRS.2016.2584107.

[11] Heming L , Qi L .Hyperspectral Imagery Classification Using Sparse Representations of Convolutional Neural Network Features [J]. Remote Sensing, 2016, 8(2): 99.DOI:10.3390/rs8020099.