Design of Semantic Analysis Model System for Spatiotemporal Information

Fan Yang[1], Zhongwang Wu[1] and Jian Xu[2,3]
[1]*Space Engineering University, Beijing, 101416, China*
[2]*Tsinghua University, 100084, China*
[3]*State Key Laboratory of Geo-Information Engineering, 710054, China*

Keywords: Text Semantic Analysis, Image Semantic Analysis, Spatiotemporal Intelligence Mining.

Abstract: There is plenty social news on the Internet, and abundant event descriptions are given in the form of text and images. The time, location, the type of events can be automatically obtained through semantic analysis of spatiotemporal information from social news. Then it is possibel to analyze the rules of events and predict the trend of events. This paper first designs a spatiotemporal intelligence semantic analysis model system, which can obtain the event type, event time and event locations, as well as the time and location rules of events based on the text semantic mining. And then the designed system can use the obtained text semantics to assist the image semantic mining to obtain the spatiotemporal intelligence such as the target type, target model, target location and action rules occurred in the event. This paper also implements the prototype system which proves that both text semantic analysis and image semantic analysis can correctly obtain spatiotemporal information.

1 INTRODUCTION

In the age of big data, we can continuously obtain the latest social news from Internet, which reports the world's trends. Through text semantic analysis and image semantic analysis, we can obtain the time, locations and types of social events from massive news, so as to analyze their rules and trends, which can be used for situation prediction. The news reports related to the same social event come from different sources, which have the forms of both text reports and images. With the progress of time, the events are also evolving. The differences between text semantics and image semantics can also assist with each other, and more accurate and rich event rules can be mined.

This paper designs a spatiotemporal intelligence semantic analysis model system, which can use the massive social news obtained from the Internet to generate the event type, time and locations, as well as the time, locations rules of the event based on the text semantic mining, and then use the obtained text semantics to assist the image semantic mining to obtain the type, model, location, time and space information such as the action rules of targets. And the prototype system is implemented in this paper, which proves that both text semantic analysis and image semantic analysis can correctly obtain spatiotemporal information.

2 RESEARCH BACKGROUND

The existing technology combining text semantics and image semantics has been applied in many fields, including image retrieval (Xie 2008, Mu et al. 2009), pathological diagnosis (Li 2009), emotion analysis (Tian 2017, Zhang 2015), and points of interest recommendation (Chen et al. 2020). Among them, references (Xie 2008, Mu et al. 2009) use the content extracted from the text semantics to retrieve the corresponding image; reference (Li 2009) comprehensively analyzes image and text semantics to obtain more accurate pathological structure and content description; references (Tian 2017) and (Zhang 2015) are both used to classify emotions by mining the semantics of Weibo Chinese text, and then use images to filter the diversity of text semantics, so as to improve the accuracy of emotion classification. Reference (Chen et. al 2020) uses the semantics of the comment text and the description of the interest points by the image semantics to comprehensively

Yang, F., Wu, Z. and Xu, J.
Design of Semantic Analysis Model System for Spatiotemporal Information.
DOI: 10.5220/0012045800003612
In *Proceedings of the 3rd International Symposium on Automation, Information and Computing (ISAIC 2022)*, pages 745-750
ISBN: 978-989-758-622-4; ISSN: 2975-9463

recommend the interest points that match user preferences.

There are also some studies (Malinowski 2021, Chaudhury et al. 2020, Singh et al. 2022, Genc et al. 2019) that detect events based on text or image data. For example, document (Malinowski 2021) converts seismic waves into images, and then extracts spatiotemporal patterns based on CNN classifier to obtain seismic events; reference (Chaudhury et al. 2020) analyzes the time feature set in motion video to determine which type of motion scene. The research on image-based event detection is also limited to the analysis of specific events, and the research on text-based event detection is relatively richer. For example, reference (Singh et al. 2022) uses a dual network called Siam network to detect and classify text data obtained in social media such as twitter, and can process data streams with a faster speed. Reference (Genc et al. 2019) is more inclined to analyze the time information in social media data to obtain time rules and cycle of detected events from appearance to disappearance.

Therefore, if the rich space-time data contained in the text and image data are comprehensively utilized, it can be used to analyze the behavior rules and action trends of targets or events. The use of text and image semantics to obtain spatiotemporal information will be conducive to dynamic tracking of social events, predicting their future trend, and timely early warning or intervention.

3 SYSTEM DESIGN

The system is to establish a semantic analysis model for spatiotemporal data analysis, including basic text processing, text semantic analysis and image semantic analysis. In order to analyze and extract the internal characteristics of spatiotemporal data, the system builds a set of semantic labeling models based on time, location and events for spatiotemporal data. The time dimension includes but is not limited to season, month, date and hour (60 minute granularity). The spatial dimension includes but is not limited to latitude and longitude and height. The event records the events that occur in the corresponding time and locations. The events recorded here need to be defined in advance.

3.1 Structure Design

The system includes the following functional modules: Web information extraction module, text extraction module, image extraction module, text data cleaning, keyword extraction module, text semantic mining module, and image semantic extraction.

Fig.1 shows the system architecture. The system can obtain web information and save it in local computers, filter local noises of web pages through the extraction of the text of web pages, and extract the text information, including plain text documents and pictures in the text. Among them, the pure text document can obtain a language that can be understood by the computer through natural language processing technology, including word segmentation, part of speech tagging and stop word filtering. The results obtained by the keyword extraction module after natural language processing can be displayed to the user's main text content, which is convenient for users to browse and process text. Through the semantic map mining module, the semantic map without annotation relationship can be obtained. After the analysis of linked data, the semantic relationship map with relation annotation can be obtained. At the same time, the system supports manual import and semantic analysis. On the other hand, the image information can obtain the type and model information of the object by using the method of image object extraction. On the other hand, the method of image semantic extraction is used to obtain the target locations and its action rules based on the results of text semantic mining.

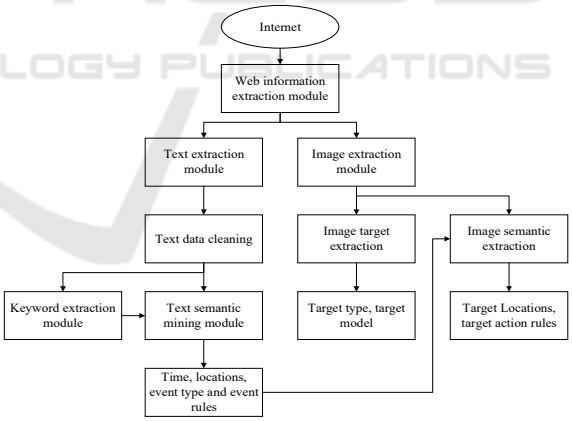

Figure 1: System architecture.

3.2 Main Functions

The system is designed to meet the teaching interaction between teachers and students, so that students can actually operate the system and understand the operation mechanism and implementation principle of the system. The system mainly has the following functions:

1) Text semantic analysis function

Text semantic analysis is mainly used to extract the entity of text content, including time dimension, space dimension and occurrence event (event type and occurrence event), and supports three-dimensional association analysis of time, locations and events. This function supports batch import of text data, model selection, and time dimension selection (season, month, date, and time). Text import supports modification and input of any text. The results are analyzed and displayed by calling the text semantic analysis model.

This function involves two operation areas: parameter input and result output. Parameter input includes batch import, model selection and time dimension selection (season, month, date and time). The result display includes: time, locations, event type and event description.

2) Image semantic analysis function

The semantic analysis of images is mainly to meet the entity extraction of batch images, including time dimension, space dimension and event (event type and occurrence event). It supports three-dimensional association analysis of time, locations and event. This function supports batch import of images, model selection, and time dimension selection (season, month, date, and time). By calling the text semantic analysis module, the results are analyzed and displayed.

This function involves two operation areas: parameter input and result output. The parameter input includes: batch import of pictures, model selection and time dimension selection (season, month, date and time). The result display includes: name, model, speed, type, time, location, longitude, dimension, event type and event description.

3.3 Implementation Principle

3.3.1 Text Data Cleaning

Before text semantic analysis, it is often necessary to clean the original text. Because the original text often contains many meaningless data, such as symbols, punctuation, or meaningless words such as "de" and "le", it is necessary to clean the useless parts. This system uses regular expressions and rules to clean the text. Among them, regular expressions are used to clear meaningless symbols and punctuation, and dictionary based word segmentation algorithm is used to clean meaningless words in the text.

(1) Regular expression (Stavros et al. 2021)

A regular expression, also known as Regex, is a sequence of characters used to match string patterns within certain text. After matching the patterns, different functions can be applied to the patterns. For example, values on a string can be replaced, and according to the regular expression patterns, values can be added or deleted in the text, and values can be searched within the text.

(2) Dictionary based word segmentation algorithm (Ling 2020)

The algorithm matches the character string to be matched with the words in an established large enough dictionary according to a certain strategy. If an entry is found, the matching is successful and the word is recognized. Dictionary based word segmentation algorithm is the most widely used and the fastest. For a long time, researchers have been optimizing based on the string matching methods, such as the maximum length setting, string storage and searching methods, and the organization structure of the vocabulary, such as using TRIE index tree and hash index.

3.3.2 Text Semantic Analysis

Text semantic analysis is mainly used to extract the entity of text content, including time dimension, location dimension and event (event type and occurrence event), and supports three-dimensional association analysis of time, locations and events. The system adopts named entity recognition algorithm (Ying et al. 2022) supplemented by rules, knowledge base and other external knowledge to realize the recognition and extraction of time, person name, institution name, location name and other spatiotemporal named entities in the text. And the event extraction algorithm (Wu et al. 2021) is used to identify and extract the event type, event trigger words, event participants and other information.

Named Entity Recognition Algorithm. One of the core tasks of this system is to effectively capture the feature information of unstructured text. Because the word segmentation task has many marked entity boundaries that are the same as the named entity task, and the corpus size of word segmentation is relatively large, the system selects the word segmentation corpus as the external knowledge and designs the character vector e^{CWS} for the word segmentation task and the character vector e^{NER} for the named entity recognition task as the input vector of the model. e^{CWS} includes external knowledge and a certain degree of noise that can provide a division basis for boundary determination of the named entity recognition task. e^{NER} can provide semantic features that are unique to the named entity recognition. The

feature representations of these two types of characters can be obtained by querying the corresponding word vector matrix.

Because the bi-directional information in a sentence is helpful for sequence modeling, which can help to judge the named entity through the above and the following, the Bi-LSTM network (Ying et al. 2022) that can capture the bi-directional information of the text is used to extract the sentence context features in this paper. Named entity recognition and word segmentation are both sequence annotation tasks, and there are strong constraints between adjacent tags. Therefore, this paper uses CRF as the decoding layer. CRF is composed of label probability matrix $E \in R^{n \times tags}$ and transition probability matrix $T \in R^{tags \times tags}$, where *n* is the number of characters in the sentence and *tags* is the number of tags.

Event Extraction Algorithm. The event extraction module extracts event information from unstructured text data, including:
1) Event trigger words: core words indicating the occurrence of events, mostly verbs or nouns;
2) Event types: ACE2005(Tan et al. 2021) defines 8 event types and 33 sub classes;
3) Event arguments: the participant of an event, mainly composed of entity, value and time;
4) Argument roles: the role of an event argument in an event.

The event extraction module uses the DMCNN (Wu et al. 2021) method to extract the event information from text data, providing a basis for text data mining. As a convolutional neural network, DMCNN also includes input layer, convolution layer and pooling layer. In the input layer, the input layer of DMCNN algorithm adopted by the system includes three types of features: CWF, PF and EF, which respectively represent word embedding, location embedding and event type embedding. The results of three embedding and splicing are used as the word level features of a word. The position embedding here actually expresses the position of each word relative to the trigger word and the candidate argument, and the event type is the type of the trigger word.

3.3.3 Image Semantic Analysis

The semantic analysis of images is mainly to meet the entity extraction of batch images, including time dimension, space dimension and occurrence event (event type and occurrence event). It supports three-dimensional association analysis of time, space and event. The system uses algorithms such as object detection and image event extraction to realize the identification of the event subject and environment, and the extraction of the contained events. The core goal of image semantic analysis is to realize the event detection module. The system adopts dual cycle multi-modal model (DRMM) (Tong et al. 2020) to realize image event detection. DRMM is used for deep interaction between images and sentences to aggregate modal features. DRMM uses pre-trained BERT and ResNet to encode sentences and images, and uses alternating double attention to select information features for mutual enhancement.

4 SPATIOTEMPORAL INFORMATION SEMANTIC ANALYSIS MODEL SYSTEM

4.1 Text Semantic Analysis Function

4.1.1 Function Design

The text semantic analysis page is divided into three areas: parameter selection area, analysis result display area and button operation, as shown in Fig. 2. The parameter selection area includes: text editing area, selection model and time dimension. The text editing area supports batch import and arbitrary text input. The time dimension supports options such as quarter, month, date and time. The system can analyze the event rules according to different time resolutions. The analysis result area includes: time, location, event type and event description. Button operations include batch import, semantic analysis and event pushing. Event pushing is to push analysis results to another system for subsequent processing by other systems. After clicking the event pushing button, the system will give a prompt of success or failure.

Figure 2: Text semantic analysis function interface.

4.1.2 Function Realization

This paper chooses the text semantic analysis model to analyze, and gets the analysis results. The text semantic analysis function can describe events in the

text, classify event types, and extract entity attributes such as time and locations.

The operation steps are as follows:

1) Click "spatiotemporal intelligence semantic analysis system" to enter the training interface, as shown in Fig. 2;

2) Click "batch import" in Fig. 2 and select the required text file to import, as shown in Fig. 3;

Model selection: for text semantic analysis model, first select time dimension by date, and then click “semantic analysis” to get the analysis result, as shown in Fig. 4.

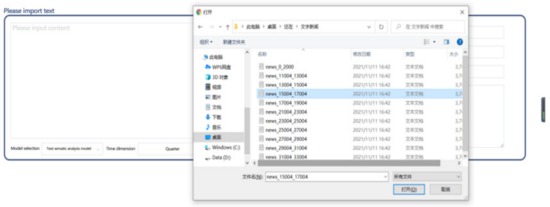

Figure 3: Select text for importing text.

Figure 4: Analysis results of text semantic analysis.

It can be seen that according to the three different pieces of news information in Fig.4, the time, locations and event type information in the text are extracted and displayed respectively, and the event rules within a period of time can be summarized and described, realizing the function of text semantic analysis.

4.2 Image Semantic Analysis Function

4.2.1 Function Design

After the text analysis, the image semantic analysis is performed using the results of the text analysis. The picture semantic analysis page is also divided into three areas: parameter selection area, analysis result display area and button operation, as shown in Fig. 5. The parameter selection area includes: image selection, model selection, and time dimensions. The image area supports batch import of images, and the time dimension supports options such as quarter, month, date and time. The system can analyze event rules according to different time resolutions. The analysis result area includes the name, model, speed, type, time, location, longitude, dimension, event type and event description of the target. Button operations include semantic analysis and event pushing. Event pushing is to push the analysis result to another system for subsequent processing by other systems. After clicking the event pushing button, the system will give a prompt of success or failure.

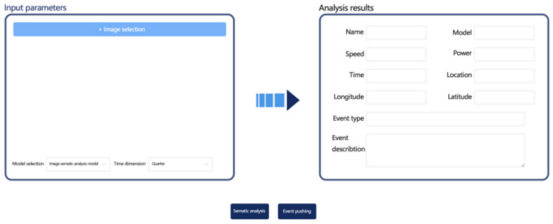

Figure 5: Image semantic analysis function interface.

4.2.2 Function Realization

The image semantic analysis function supports importing image files, selecting an image semantic analysis model for analysis, classifying the objects in the image, determining the model, determining the target location, and displaying the event type and event description results. The correspondence between images and text is n:1, that is, there is one text file extracted from a news release, and there are N pictures extracted from the same news release. Therefore, when multiple images are selected, the text files corresponding to multiple images are also selected, and the locations, time and event information obtained by semantic analysis are also displayed in the results.

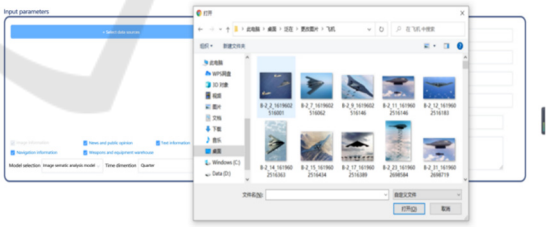

Figure 6: Select images for importing image data.

Figure 7: Select model and time dimension to obtain analysis results.

The operation steps are as follows:

1) Click "spatiotemporal intelligence semantic analysis system" to enter the training interface;

2) Click "image selection" in Fig. 5 and select the required picture to import, as shown in Fig.6;

3) Model selection: for image semantic analysis model, first to select time dimension by quarter, and then click semantic analysis to get the analysis result, as shown in Fig. 7.

It can be found that the type and model of the target are obtained according to the multiple images, so as to match the speed, power and other intelligence information of the target from the database. Using the results of text semantic analysis, the type judgment and event description of events, such as ship events, can be obtained at the same time.

5 CONCLUSIONS

This paper designs a spatiotemporal intelligence semantic analysis model system, which can extract images and text information from the massive news events obtained from the Internet, and conduct semantic analysis on the texts and images respectively to obtain the information about the time, locations, types and rules of the events. The system (1) supports the management of the text through the quality requirements of the text data, retains the text with analysis value, and removes the dirty data; (2) It supports the establishment of semantic analysis model and the extraction of text content, including time dimension, space dimension and occurrence event (event type and occurrence event); (3) It supports time information to season, month, date, time-sharing granularity (60 minutes, etc.), and analyzes the intrinsic value of information in the time dimension; (4) It supports the use of events (event types and events) to classify texts, analyzes the change rules of similar events in the two dimensions of time and locations, mines the potential characteristics of events, and provides guidance for future decision-making. With the generated spatiotemporal information and spatiotemporal movement rules, it is possible for us to make predictions on target intention as our future work.

REFERENCES

Chaudhury S, Kimura D, Vinayavekhin P, et al. Unsupervised Temporal Feature Aggregation for Event Detection in Unstructured Sports Videos[J]. 2020.

Chen Jianbing, Shen Jianfang, Chen Pinghua, Point of Interest Recommendation Integrating Review and Image Semantic Information [J], *Computer Engineering and Applications*, 2020, 56(19): 160-167.

Genc, H., Yilmaz, B. (2019). Text-Based Event Detection: Deciphering Date Information Using Graph Embeddings. In: Ordonez, C., Song, IY., Anderst-Kotsis, G., Tjoa, A., Khalil, I. (eds) Big Data Analytics and Knowledge Discovery. DaWaK 2019. *Lecture Notes in Computer Science*, vol 11708. Springer, Cham. https://doi.org/10.1007/978-3-030-27520-4_19.

Li Bo, Analysis Model of Medical Text and Image based on LDA and LSA and its Application [D], *Jilin University*, 2012.

Ling Zhao, Ailian Zhang, Ying Liu, Hao Fei, "Encoding multi-granularity structural information for joint Chinese word segmentation and POS tagging", *Pattern Recognition Letters*, 138: 163-169, 2020.

Malinowski M. Automatic Image-Based Event Detection for Large-N Seismic Arrays Using a Convolutional Neural Network[J]. *Remote Sensing*, 2021, 13.

Mu Yakun, Feng Shengwei, Zhang Jin, Image Retrieval Based on Text and Sematic Relevance Analysis [J], *Computer Engineering and Applications*, 2009, 55(1):196-202.

Singh T , Kumari M , Gupta D S . Real-time event detection and classification in social text steam using embedding[J]. *Cluster Computing,* 2022:1-19.

Stavros Konstantinidis, Nelma, Moreira, Rogério Reis, Partial derivatives of regular expressions over alphabet-invariant and user-defined labels, *Theoretical Computer Science*, Volume 870, 16 May 2021, Pages 103-120.

Xiaokao Tan, Guofeng Deng, Xiangjun Hu, Multi-granularity context semantic fusion model for Chinese event detection, *ICICSE 2021: 2021 10th International Conference on Internet Computing for Science and Engineering*, July 2021, pp: 1–7.

Tan Junxin, Research on Sentiment Classification for Microblogging based on Multimodel Data [D], *Nanjing University*, 2017.

Tong M , Wang S , Cao Y , et al. Image Enhanced Event Detection in News Articles[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(5):9040-9047.

WU Fan, ZHU Peipei, WANG Zhongqing, LI Peifeng, ZHU Qiaoming, Chinese Event Detection with Joint Representation of Characters and Word, *Computer Science*, 48(4), 2021.

Xie Lin, Integrating Textural Semantic and Visual Content for Web Personal Image Retrieval [D], *Beijing Jiaotong University*, 2008.

Ying An, Xianyun Xia, Xianlai Chen, Fang-Xiang Wu, Jianxin Wang, "Chinese clinical named entity recognition via multi-head self-attention based BiLSTM-CRF", *Artificial Intelligence In Medicine*, 127: 102282, 2022.

Zhang Yaowen, Research on Sentiment Classification for Microblogging with Text and Image [D], *Nanjing University,* 2015.