

Research Article

A Cognition Knowledge Representation Model Based on Multidimensional Heterogeneous Data

Dong Zhong ¹, Yi-An Zhu,¹ Lanqing Wang ¹, Junhua Duan,¹ and Jiaxuan He²

¹School of Computer Science, University of Northwestern Polytechnical of China, Xi'an, Shaanxi 710129, China

²School of Software, University of Northwestern Polytechnical of China, Xi'an, Shaanxi 710129, China

Correspondence should be addressed to Dong Zhong; dzhong@nwpu.edu.cn

Received 26 September 2020; Revised 26 November 2020; Accepted 14 December 2020; Published 29 December 2020

Academic Editor: Fei Xiong

Copyright © 2020 Dong Zhong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The information in the working environment of industrial Internet is characterized by diversity, semantics, hierarchy, and relevance. However, the existing representation methods of environmental information mostly emphasize the concepts and relationships in the environment and have an insufficient understanding of the items and relationships at the instance level. There are also some problems such as low visualization of knowledge representation, poor human-machine interaction ability, insufficient knowledge reasoning ability, and slow knowledge search speed, which cannot meet the needs of intelligent and personalized service. Based on this, this paper designs a cognitive information representation model based on a knowledge graph, which combines the perceptual information of industrial robot ontology with semantic description information such as functional attributes obtained from the Internet to form a structured and logically reasoned cognitive knowledge graph including perception layer and cognition layer. Aiming at the problem that the data sources of the knowledge base for constructing the cognitive knowledge graph are wide and heterogeneous, and there are entity semantic differences and knowledge system differences among different data sources, a multimodal entity semantic fusion model based on vector features and a system fusion framework based on HowNet are designed, and the environment description information such as object semantics, attributes, relations, spatial location, and context acquired by industrial robots and their own state information are unified and standardized. The automatic representation of robot perceived information is realized, and the universality, systematicness, and intuition of robot cognitive information representation are enhanced, so that the cognition reasoning ability and knowledge retrieval efficiency of robots in the industrial Internet environment can be effectively improved.

1. Introduction

Human beings understand the environment and recognize the objects in the environment according to the shape and color of the objects, and the information of environment and objects acquired by human beings is stored in the human brain in a structured and hierarchical form. With the improvement of computing capacity and the amount of stored data of computing system, the machine can construct the environment knowledge base by simulating human cognition and storing information and then has the ability to recognize the environment like human beings.

In the working environment of industrial Internet, describing the service environment information and storing it in the form of knowledge base is the premise of realizing

intelligent and personalized service. When robots perform tasks cooperatively in the industrial Internet, the characteristics of diversity, semantic, hierarchy, and relevance of the involved environmental information greatly increase the difficulty for robots to complete intelligent service tasks. By characterizing the environment information and storing it in the form of knowledge base, it can help robots not only master the semantic information of objects in the working environment but also help robots search for objects in the environment quickly and improve the ability of intelligent cooperative work of robots. However, in the field of robot knowledge base construction, there are some problems in the aspects of environment representation and environment information acquisition, such as insufficient practicability, poor human-machine interaction capability, insufficient

knowledge base expansion, high degree of manual participation in semantic information addition, and low reliability of independent information addition.

Knowledge graph technology, which has arisen in recent years, can effectively solve the above problems. The knowledge graph was put forward by Google in 2012 and used to improve search engines. Subsequently, many scholars began to study it one after another, and many achievements emerged in the field of knowledge storage and inquiry. A large number of structured knowledge data are usually stored in knowledge graphs, and the real-world scenes are directly modeled by using triplets being made of nodes and relationships. Using the universal “language” of triplets, the relationship between objects and objects in the scene can be effectively and intuitively represented. In addition, the knowledge graph emphasizes not only the concept and but also the entity relation and the entity attribute value, which enriches and expands the knowledge of ontology. As a graph-based data storage structure, knowledge graph based on the above characteristics not only has strong storage, query, and reasoning capabilities but also has advantages in real-time update and human-machine interaction.

The application of knowledge graph technology to the construction of robot knowledge base in industrial Internet can greatly improve the ability of the robot to represent and store environmental knowledge. It can effectively unify various environmental information into context information that the robot can understand. At the same time, the knowledge base based on knowledge graph stores environmental information in the form of structured network, which makes the robot knowledge base query have similar associative ability to human beings and becomes the key to improve the robot’s intelligence and realize service tasks.

2. Related Work

Environmental information representation and knowledge base construction are the research hotspots in the field of industrial Internet and are the key steps for robots to realize intelligent services in industrial Internet. The essential problem lies in how to encode and store the environment information in a unified form, and based on this encoding, robots can perceive and understand the environment and then realize intelligent cooperative work. Therefore, the process of building knowledge base is divided into two steps: acquiring environment knowledge and representing and storing knowledge.

Considerable achievements have been made in obtaining environmental information. Kollar et al. [1] proposed a task-based human-machine dialogue learning method for environmental knowledge and analyzed the entity mapping from each semantic element to the environment by introducing a joint probability distribution model on semantics. Lemaignan et al. [2] proposed a dialogue module to convert the natural language into symbolic facts (OWL statements) according to the intention of the original sentence and constructed a symbolic knowledge base. Lemaignan et al. [3] established a basic and shared world model by extracting,

representing, and using symbolic knowledge in the verbal and nonverbal interaction between humans and robots, which is suitable for subsequent high-level tasks, such as dialogue understanding. Schiffé et al. [4] proposed a flexible natural language interpretation system for mobile robot voice instructions, modeling language processing as an interpretation process that maps speech to entities in the environment. In addition, multimedia interaction and autonomous perception of robots are also important ways to acquire knowledge. Burghart et al. [5] proposed a two-layer embedded framework for service robots, which gradually embedded perception, learning, action planning, motion control, and human-like communication into the architecture. Mozos et al. [6] use robots to learn examples of common models of furniture on the Web. Through these models, the classification and location of unknown furniture in the environment are realized. Ko et al. [7] proposed semantic map representation and human-like navigation strategies for monocular robots, extracting semantic information on the basis of constructing visual maps and constructing semantic maps represented by nodes such as regions or landmarks and their spatial relationships. Song [8] embeds environmental information into artificial signposts, by identifying which robots can acquire environmental information. Tamas and Cosmin Goron [9] proposed a three-dimensional point cloud marking system based on physical characteristics, which can realize semantic interpretation of surrounding environment items (such as furniture, ceiling, and door). Literature studies [10, 11] construct an intelligent space suitable for robot task execution by adding robot-perceivable tags in the environment. The intelligent space distributes the semantic information needed by the service in the surrounding environment, and the robot obtains the corresponding information through corresponding sensing methods to complete the service task.

With the development of robot technology and artificial intelligence technology, more and more attention has been paid to the human-machine interaction of environmental information. At present, the commonly used representation methods of environmental information mainly include predicate logic representation, production rule representation, and semantic Web ontology representation, etc. Ontology’s structured knowledge representation has a wide range of representation, strong representation ability, and reasoning ability, which has quickly become a research hotspot. Literature studies [12–15], respectively, use ontology technology to construct corresponding knowledge frameworks to realize environmental representation. Park et al. [12] proposed a scene knowledge base system design based on domain knowledge ontology. Hao et al. [13] used the ontology knowledge base model to reorganize the original dataset, making the logical structure of the new dataset more suitable for upper application and improving the utilization rate of open data. Yang et al. [14] proposed a semiautomatic annotation framework which was represented by WoT resource metadata. The framework is based on a probabilistic graphical model, mapping the schematic of WoT resources to independent domain knowledge base and collectively inferring entities, classes, and relationships.

Das et al. [15] designed an ontology-based information sharing mechanism among robots, forming a collective knowledge base, which is convenient for the overall control and planning of the system. In addition to ontology representation, Jezek Moucek [16] designed a semantic framework and realized object-oriented environmental representation through semantic Web language. Chen [17] proposed an environment representation method based on quadtree. By designing an access code mechanism, the robot can quickly grasp the environment obstacle information to complete navigation tasks in complex scenes. Gao et al. [18] designed a three-layer representation model of the indoor environment, which represented the family environment in the form of holographic map and applied it to object-oriented task service.

These achievements not only reflect the progress in the research direction of environment information representation but also show some problems, such as lack of structure and expansibility of knowledge representation in the aspect of environment representation; the acquisition of representation information is characterized by high manual participation and low information reliability. Therefore, it is urgent to study the information representation method and representation information acquisition methods suitable for the intelligent service environment of industrial Internet.

The concept of knowledge graph was put forward by Google in 2012 [19] to improve the performance of its search engine. As soon as it was put forward, it aroused a strong response and got a widespread concern. Many organizations followed up quickly and carried out related research in succession to improve the performance of their own search engines. Many achievements have emerged, such as Sogou's knowledge cube and Baidu's bosom friend [20, 21]. Since then, knowledge graph technology has gradually spread to other fields.

In order to improve the accuracy of search, Marino et al. [22] combine the knowledge graph with neural network and introduce the graphic search neural network method to effectively merge the large knowledge graph into the visual classification pipeline to improve the accuracy of image classification with structured prior knowledge. Szekely et al. [23] proposed a method of constructing knowledge graph, which uses semantic technology to fuse data from different sources, extends crawled data to billions of triplets, and applies knowledge graph to the DIG system to combat human trafficking. Literature studies [24–29], respectively, constructed knowledge graphs of corresponding disciplines and carried out visual analysis. Among them, GDM Laboratory of Fudan University designed the Chinese knowledge graph related to books to realize the visualization of the classified query in the book field [27]. Jia et al. [29] constructed TCM (traditional Chinese medicine) knowledge graph, realized effective integration of TCM knowledge resources, and discussed the application prospect of TCM knowledge graph. Lu et al. [30] constructed the teaching knowledge map, discussed the knowledge graph with students, and compared the two, which may enable the teacher to reflect on their own work and cause the students to realize questions related to study, and thus, it is advantageous to enhance the teaching content, strategy, and activity.

In order to construct the corresponding database, Kumar et al. [31] use the structured data based on knowledge graph to construct a unified system combining speech recognition and language understanding, which is used to solve the specific problems of automatic speech recognition and language processing. Jia et al. [32] proposed a network security knowledge base and inference rules based on a five-element model. Using machine learning, entities are extracted, ontologies are constructed, and network security knowledge base is obtained. Kem et al. [33] proposed a knowledge model to describe the spatial structure of an environment. The network physics society entities it contains, and the relationship between them, are regarded as a graph, called cyberspatial graph (CSG). Collarana [34] designed a semantic data integration method called FuhSen, which uses keywords of Web data sources and structured search function to generate knowledge graphs to merge data collected from available Web data sources. Kumar et al. used a semantic-rich knowledge graph library to solve specific problems in automatic speech recognition and natural language processing.

These achievements show that the knowledge graph has a very broad development prospect, but most of the achievements and applications exist in search engine improvement, text information processing, and intelligent search. Hao et al. [35] proposed an environmental information representation mechanism based on knowledge graph in the field of intelligent service of home robots. In this paper, we further into the professional database in the Internet industry for the construction of knowledge graph based on the above research, at the same time in the process of knowledge extraction added the description information of work environment, the function information, and real-time status information of the mechanical arm into knowledge extraction triples, and make the mechanical arm cognitive external working environment and working object. And the mechanical arm can match the external environment and the object information to its own state and ability, so that it has the ability to judge whether it can complete some nonspecific complex tasks.

3. Machine Cognition Knowledge Graph Framework

In the environment of industrial Internet, robots can obtain the information of the robot's external environment and its internal state through various sensors and interaction mechanisms, which are collectively referred to as robot perception information in this paper. The expression and storage of perceptual information is the premise for robots to realize cognitive ability. However, when robots complete nonspecific complexity tasks, it is far from enough to simply recognize objects and perceive scenes. In order to achieve higher intelligence and autonomy, robots need to master not only various attributes of objects operated, such as location, function, operation mode, and instructions but also the categories of objects. In this section, aiming at the representation of robot perception information, considering the diversity, semantics, hierarchy, and relevance of the

representation of perception information when robots perform nonspecific complexity tasks, a robot cognitive knowledge graph framework including perception layer and cognition layer based on knowledge graph is designed to realize the unified representation and storage of semantic information, attribute information, spatial function information, and context information of operating objects in the environment, thus realizing the construction of robot cognitive knowledge base.

The knowledge representation method based on semantic Web, represented by knowledge graph, not only has the advantage that predicate logic representation is easy to understand, but also knowledge always exists in the form of triplets (entities, relationships, and entities) in the knowledge graph, which can represent knowledge in a formal and concise way and intuitively model various scenes in the real world. In the knowledge graph, the entities of the triplets are regarded as nodes and the relationships between the entities are drawn as edges, so the knowledge base containing a large number of triplets forms a huge knowledge network. The method can effectively solve the problem that the robot is difficult to learn and store the relationship between objects in complex working scenes and improve the robot's ability in human-machine interaction, object inquiry, knowledge reasoning, and task execution.

Knowledge base aims to describe various entities or concepts and their relationships in the real world. Knowledge base is modeled as a set of triplets, with entities represented by e and relationships represented by r . In the knowledge graph, entities or concepts are represented by nodes and attributes or relationships are represented by edges, and thus, entities and relationships in the real world can be formed into a huge semantic network graph. The following gives the definitions of three kinds of nodes and edges included in the robot cognitive knowledge graph.

Entity. It refers to instances (people or objects, etc.) in the environment, such as a person, a table, and a gear. Entities are the most basic elements in the knowledge graph. Each entity has a unique number in the knowledge base to distinguish it from other entities.

Concept. It is a collection of entities with the same characteristics, for example, gears and bearings are both workpieces machined, and lathes and milling machines are both production equipment.

Attribute. It is used to describe certain characteristics of an entity, for example, the shape parameters and performance parameters of gears or bearings. In the cognitive knowledge graph, it exists in the form of edges, and attributes are measured by attribute values.

Relationship. It is formalized as a function, and it exists as an edge in the cognitive knowledge graph and is used to describe the relationship between entities or concepts in the graph.

Based on the above definitions, triplets are an information representation of cognitive knowledge graphs,

namely, $G = (E, R, S)$, where $E = \{e_1, e_2, \dots, e_n\}$ is an entity set and $R = \{r_1, r_2, \dots, r_n\}$ is a relational set. The ternary combination set of representation information S is $S \subseteq E \times R \times E$.

The construction of the cognitive knowledge base based on the cognitive knowledge graph is divided into the construction of the perception layer and the construction of the cognition layer. The flowchart is shown in Figure 1. The perception layer is mainly composed of a series of factual data, including a series of multisource heterogeneous original data and coded structured data. The original data include environmental data collected by robots through various sensors, data acquired by robots through a universal knowledge base on the Internet, and professional data stored in some professional databases. The cognition layer is mainly composed of a series of knowledge stored in units of triplets, such as triplets with gears in the accessory cabinet (gears, stored in, accessory cabinet). The cognition layer is built on the perception layer and is the core of the knowledge graph, storing abstract knowledge (concepts).

3.1. Formation of Perception Layer. The construction of the perception layer needs to firstly discover and add entities (i.e., find entities), so the data source of knowledge graph is the first problem faced by the self-construction of knowledge graph. According to the working environment of the industrial Internet, the robot can sense and collect images, videos, point clouds, and other data from the environment by carrying corresponding sensors for environmental object identification and space division, etc. However, a lot of tacit knowledge (such as functional attributes and operational attributes of objects) is difficult for the robot ontology to acquire through sensors. When people ask the robot to complete a nonspecific and complicated task of "opening the door," the robot neither knows the functional attribute of the key to open the door, nor does it know the operating attribute of the key to correctly insert the key into the lock hole, rotate twice, and then turn the doorknob counterclockwise by 90 degrees, so the task cannot be completed. Robots learn relevant knowledge from encyclopedia websites (e.g., Wikipedia), electronic product manuals, and shared structured data sources (e.g., relational databases of various specialties) and add them to the knowledge base, which is an effective way to solve the above problems. Due to the different data sources of the knowledge base, we have designed a knowledge extraction module to achieve the consistency of data formats. The structure diagram is shown in Figure 2, which mainly includes a triplet downloader, a document downloader, an entity and entity relation processing submodule, a triplet filter, a robot capability description module, and an environment information extraction module. The main functions are as follows:

- (1) The document downloader is designed based on crawler technology. Its main work is to capture the text of the web page and other text knowledge provided by the Internet. The robot grabs the text of the web page through the document downloader and downloads it to the local and processes the text of the

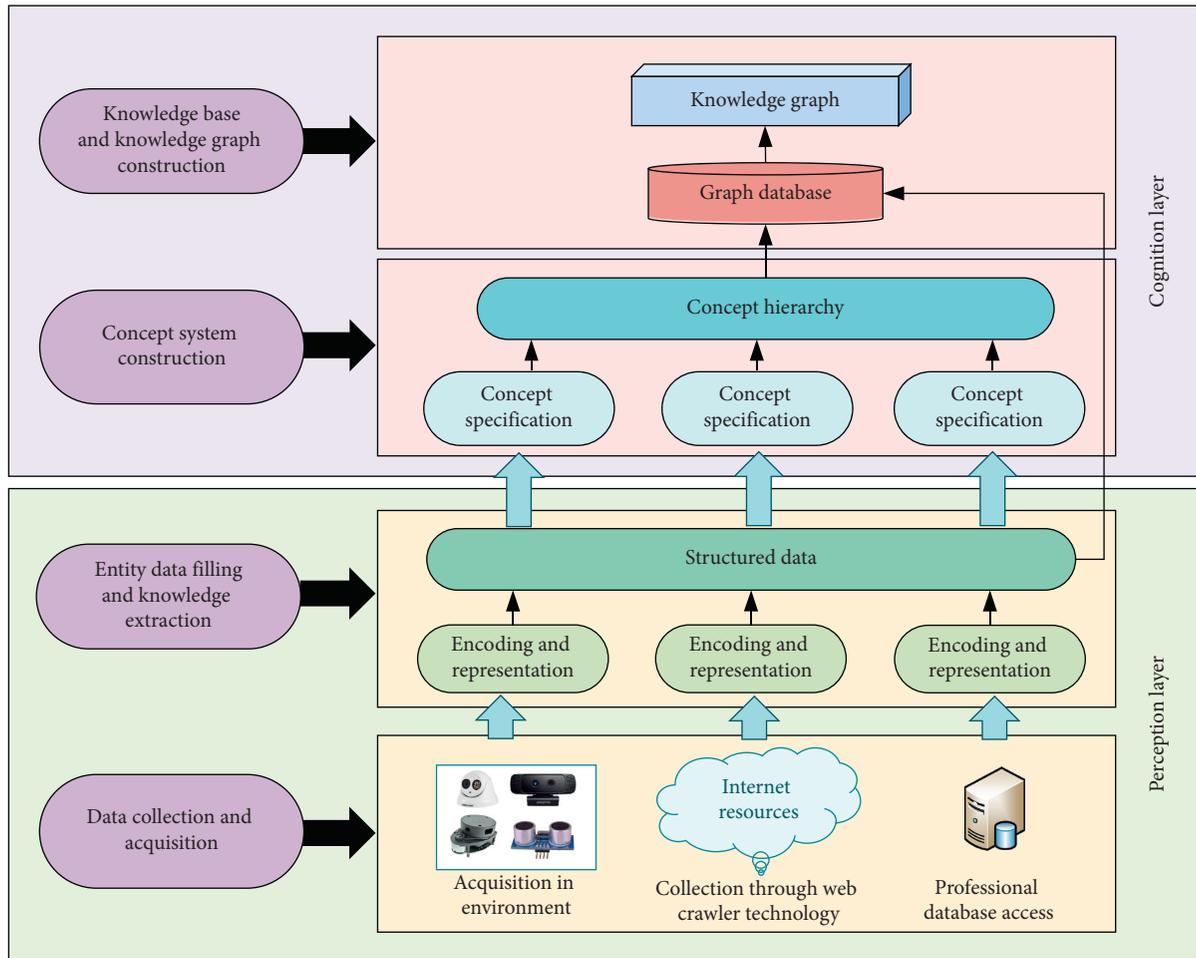


FIGURE 1: Flowchart of knowledge graph construction.

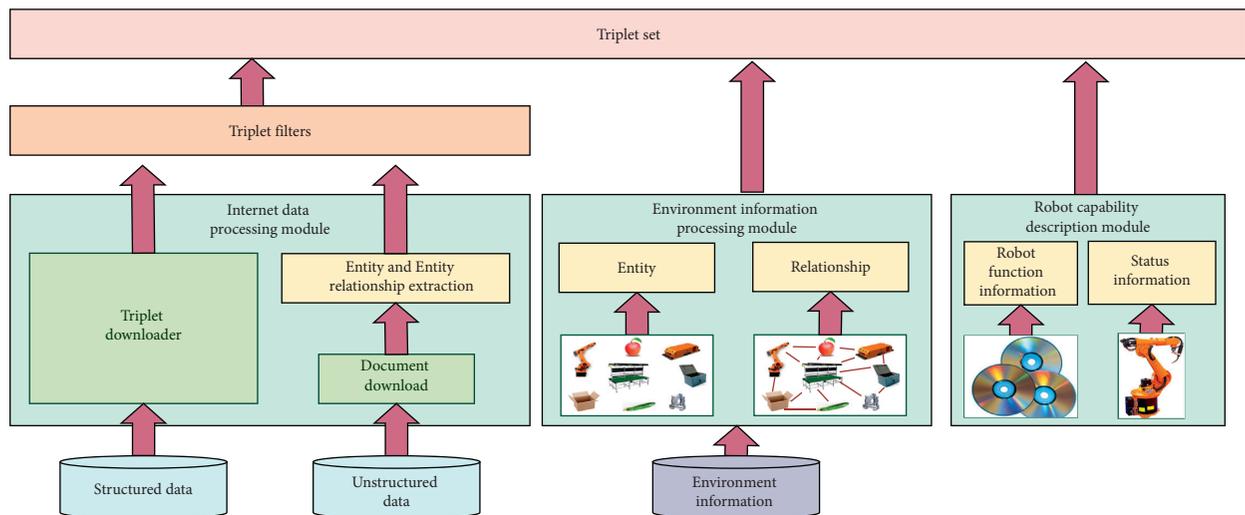


FIGURE 2: Logical structure of knowledge extraction.

web page to remove invalid data, thus obtaining the text data to be learned.

(2) Entity and entity relationship extraction module: it is based on the open-source toolkit CORENLP

developed by Stanford University for natural language processing. Through the use of the NER analysis module in CORENLP, lexical features of statements are analyzed to realize automatic

extraction of “text” entity relationships obtained by the above document downloader.

- (3) For structured data, there is no need to extract triplets through entity and entity relation module, and we can directly download triplets from the data source through the triplet downloader as candidate triplets to be added to the triplet candidate set.
- (4) Triplet filters: entities and entity relationships extracted by robots from unstructured and structured data inevitably contain repeated information. Through the triple filter module, repeated addition of information is avoided.
- (5) Environment information processing module: it obtains the location and attribution relationship between entities and entities in the environment through semantic SLAM technology [36] and spatial structured reasoning technology [37], respectively, for information perceived by the robot environment.
- (6) Robot capability description module: for the capability description information of the robot, the semantic description information of the capability of the robot in different states and the capability of operating different entities in the current environment is obtained through the description information of the robot and the current state information. For example, what objects can be operated on in the current environment and what operations can be performed on the objects under what posture.

Through the knowledge extraction module, the robot extracts entities and relationships from the working environment, its own state information, and unstructured and structured data and then generates triplets. However, the triplet set at this time is not the final knowledge graph, and entity disambiguation and entity alignment are needed for triplets in the set to realize the fusion of heterogeneous knowledge from different sources.

3.2. The Formation of Cognition Layer. Knowledge graph technology, like ontology technology, uses the triple set of nodes and relationships to directly model real-world scenes. Using the universal “language” triple set, we can effectively and intuitively represent the relationship between objects in the scene. In terms of content, compared with ontology, which emphasizes conceptual relationship, knowledge graph emphasizes not only the concept but also the entity relationship and the entity attribute value, which enriches and expands ontology knowledge as shown in Figure 3.

Therefore, the construction concepts in robot cognitive knowledge graph are similar to ontology, which is formed by identifying, classifying, and abstracting entities in the environment; we establish the up-down relationship between concepts according to the hierarchy and connection between concepts, thereby forming a conceptual hierarchy as shown in Figure 4. Therefore, the construction of the cognition layer in the cognitive knowledge graph is divided into two steps: the extraction of concepts and the establishment of the relationship between concepts.

3.2.1. The Extraction of Concepts. At present, most of the research studies on concept extraction in knowledge graph focus on the statistics of text corpus and extract candidate concepts in a certain field by calculating the frequency of words in various documents. This paper also uses statistical methods to extract candidate concepts in the working environment of industrial Internet. According to experts’ experience, a concept is a word or phrase that appears at a high frequency in the field, which can represent the characteristics of the field. Therefore, the concept of specific working environment of industrial Internet can be defined by the following two features: (1) the frequency of occurrence in specific application fields of industrial Internet is higher than that in other fields and (2) it is distributed evenly in the documents of specific application fields of industrial Internet, rather than concentrated in the documents of individual industrial Internet application fields.

For the abovementioned feature (1), we quantitatively describe it with domain relevance. Domain relevance, as its name implies, is used to describe the appropriateness of concepts and domains, and its calculation formula is as follows:

$$\text{Domain } R_{n,m} = \frac{P(n/D_m)}{\max P(n/D_i)}, \quad 1 < i < t. \quad (1)$$

Set of domain = $\{D_1, D_2, \dots, D_m\}$, and formula (2) shows the calculation of the correlation between concept n and domain. The conditional probability $P(n/D_m)$ in the formula can be estimated by the following formula:

$$E\left(P\left(\frac{n}{D_m}\right)\right) = \frac{f_{n,m}}{\sum_{n' \in D_m} f_{n',m}}, \quad (2)$$

where $f_{n,m}$ is the frequency of concept n in domain D_m and $f_{n',m}$ is the frequency of concept n' in D_m . As shown in formula (2), the domain correlation of a concept is only proportional to its frequency in the domain and has nothing to do with other elements. When a certain concept frequently appears in individual domain documents, it will be ineffective to measure the concept only by using domain relevance, so feature (2) is needed to reflect the distribution of concepts in domain documents.

For feature (2), we use domain consistency to quantitatively describe it, and its calculation formula is as follows:

$$\begin{aligned} \text{Domain } C(n, D_m) &= H(P(n, d_j)) \\ &= \sum_{d_j \in D_m} \left(P(n, d_j) \log \frac{1}{P(n, d_j)} \right), \\ E(P(n, d_j)) &= \frac{f_{n,d_j}}{\sum_{d_j \in D_m} f_{n,d_j}}, \\ W_{n,m} &= \alpha \text{Domain } R_{n,m} + \beta \text{Domain } C_{n,m}, \quad \alpha, \beta \in (0, 1), \end{aligned} \quad (3)$$

where d_j is any document in the field, f_n , and d_j is the frequency of concept n appearing in document d_j . Formula (3) combines domain relevance and domain consistency,

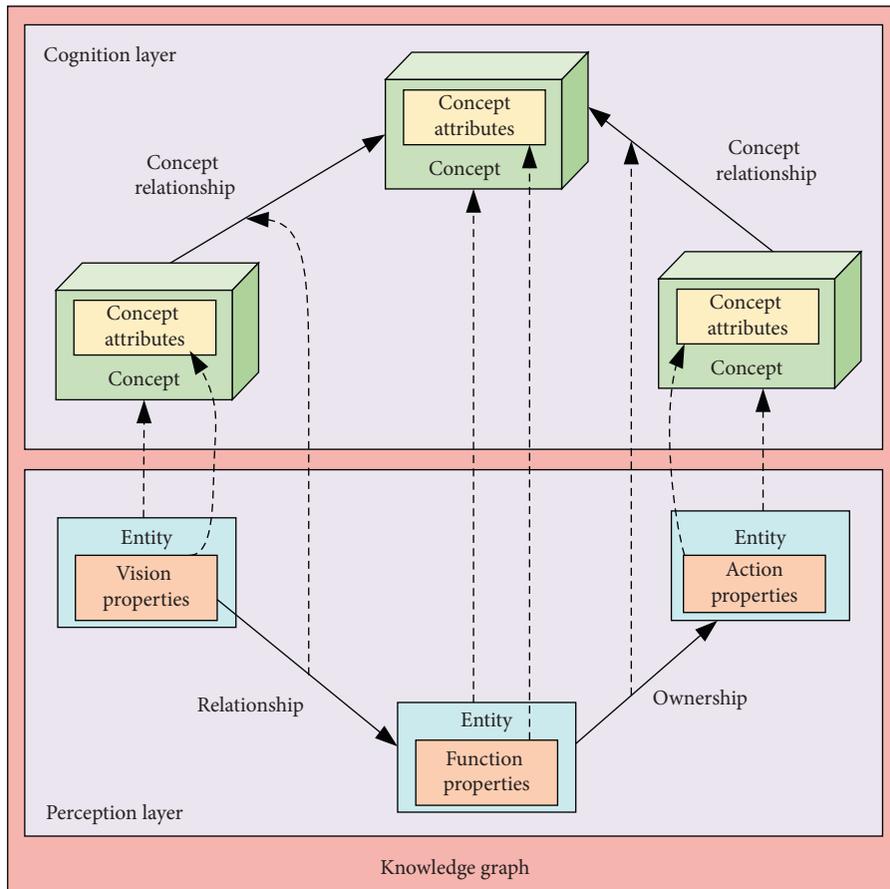


FIGURE 3: The relationship of knowledge graph and ontology.

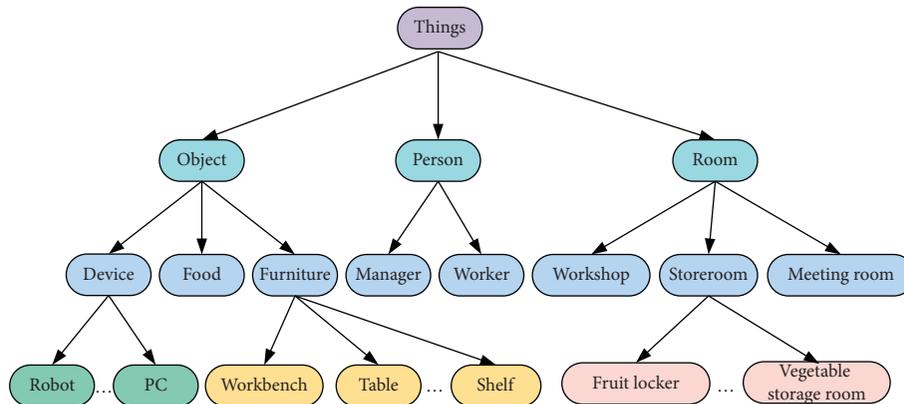


FIGURE 4: Hierarchical relationship of concepts.

and by setting related parameters α and β , the probability of concepts unrelated occurrence to the domain can be reduced.

3.2.2. *The Establishment of the Relationship between Concepts.* In the previous section, we extracted the related concepts of specific application fields of industrial Internet, but these concepts are discrete and cannot be linked with each other. In the knowledge graph, concepts are connected by the

relationship between upper and lower levels, so we need to establish the relationship between upper and lower levels of these concepts. In people’s language usage habits, there are often such sentence patterns as “A is a B,” “A is like B and C,” and “A has B, C”. Using this kind of language model, we can establish the hyponymy relationship between A, B, and C. For example, “unmanned factories have production equipment such as lathes and milling machines”. According to the above linguistic schema, it can be deduced that the upper words of lathes and milling machines are production equipment.

Using the above linguistic sentence patterns, robots in industrial Internet can learn the upper and lower relations of concepts in encyclopedic text resources related to their specific application fields, thus connecting discrete concepts in series with the upper and lower relations to form a structured and hierarchical concept layer.

As shown in Figure 5, the robot combines the perceptual information perceived by the robot with the semantic description information such as attributes, concepts, and relationships of entities obtained from the Internet and establishes a structured reasoning knowledge base including the perception layer and the cognition layer, thereby realizing the unified representation and storage of the semantic information, attribute information, spatial function information, and context information of objects in the environment and generating the corresponding cognitive knowledge graph.

4. Knowledge Fusion Algorithm Based on Vector Features

In the working environment of industrial Internet, when completing some unspecified complex tasks, robots usually need to know a variety of attributes of operated objects, such as position, function, and operation. The knowledge base with good performance and detailed content is the basis to solve this problem. In the process of building knowledge base, knowledge extraction technology realizes data acquisition. However, due to its wide data sources, the structure of these data is also very different; due to the freedom of natural language expression, polysemy often occurs. The robots based on this kind of knowledge base cannot generate unified and unambiguous cognition or complete some unspecified and complex tasks, so it is necessary to fuse the extracted knowledge.

In this paper, the differences of entity semantics and knowledge system in robot knowledge base are studied. To solve the problem of entity semantics inconsistent, this paper proposes an entity semantic fusion method, which calculates its word vectors through text information and description information of structured knowledge and determines the similarity of entity semantics according to the cosine similarity of word vectors, thus realizing the alignment and disambiguation of entity semantics. For the differences of knowledge systems, a framework of knowledge system integration based on HowNet is designed, which effectively solves the problem of knowledge system integration.

4.1. Semantic Fusion of Multimodal Entities Based on Vector Features. For the fusion of heterogeneous knowledge, a cross-modal data representation is needed first. In recent years, it is becoming more and more popular to represent various types of data in the form of vector features and to conduct corresponding research based on them. The most common vector feature learning is text domain, such as text embedded representation or distributed word representation. They are based on unsupervised learning and only rely on the input text corpus. The heterogeneous text entity

information is encoded into dense vector representation, and the similarity between two text entities is obtained by calculating the cosine similarity between entity vectors, thus completing the entity fusion.

Based on this idea, some structured data can also be encoded in low-dimensional vector space. For example, knowledge graph or ontology database (DBpedia and Wikidata), which is a network database, composed of entity-relation-entity or entity-predicate-entity. After the structured data composed of these entities, relations, or predicates are coded into low-dimensional vectors, it is easier to predict the links in the knowledge graph and reason the knowledge.

Therefore, it is a practical and effective method to preprocess text knowledge and structured knowledge by corresponding means and represent them in vector form and then realize multisource heterogeneous knowledge fusion by vector fusion.

4.2. Vectorized Representation of Text Information. For the vector transformation of text corpus knowledge, we choose word2vec model. word2vec is an open-source tool for word vector calculation released by Google in 2013, which has attracted the attention of industry and academia and become a popular application in the field of natural language.

Figure 6 shows the flow of generating word vectors based on word2vec method and obtains input and output word pairs by Skip-gram or CBOW (Continuous Bag of Words) model. At the same time, the input and output words are encoded in one-hot mode, which is taken as a training sample and brought into a neural network for training. The input matrix is multiplied by the output-hidden layer weight matrix, and the result is the word vector of the output word.

CBOW and Skip-gram are two different methods of word2vec. CBOW predicts the probability of a word according to its context. Skip-gram is just the opposite. In the process of solving word vectors, these two models actually play an important role in obtaining training samples needed for subsequent neural network model training.

Taking Skip-gram algorithm as an example, there are two parameters in this method: skip-window and num-skips, which, respectively, represent the number of contexts to be predicted and the number of output results. Taking the sentence "Robots walk around boxes to the workplace" as an example, assuming that the input word is box, skip-window and num-skips take 2 and 4, respectively, the final words containing box are (walk, around, box, to, the), and the result can be represented in the form of (input word, output word). Input all the words in the sentence as input words, and the output results are shown in Table 1.

The training samples as shown in Table 1 are obtained through the skip-gram model, but these training samples cannot be directly used for neural network training. One-hot coding is also needed to encode them to mark the position where words appear as 1 and the rest as 0. For example, the code of robots is [1, 0, 0, 0, 0, 0, 0] in the above example.

Enter the one-hot code of the input words into the neural network model as shown in Figure 7, and each neuron in the

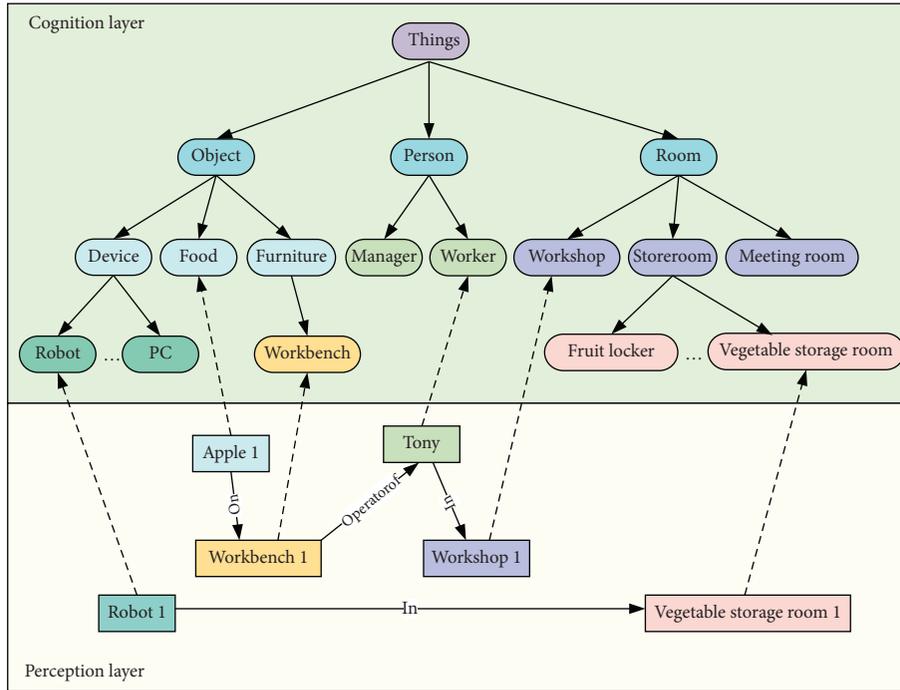


FIGURE 5: Schematic diagram of the cognition layer and perception layer in knowledge base.

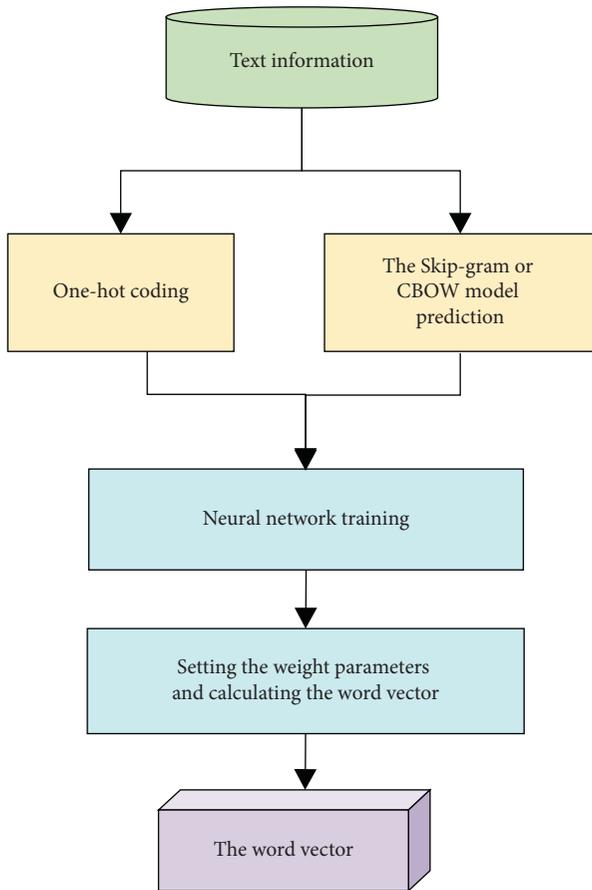


FIGURE 6: Flowchart of word vector formation.

input layer represents each bit in one-hot code. A hidden layer neural network is set up, and the one-hot coding passes through an input-hidden layer weight matrix and a hidden layer-output matrix to finally obtain the output of the neural network. We do not care about the output of the neural network, but the input-hidden layer parameter weight matrix after neural network training. The corresponding word vector can be obtained by multiplying the one-hot code of words with the parameter matrix.

4.3. *Vector Representation of Structured Knowledge.* TransE model is an algorithm to transform structured knowledge into vector features, which was proposed by Bordes [38] in 2013. Figure 8 shows a schematic diagram of the distributed vector representation model based on entities and relationships. Figure 9 shows the algorithm. The intuitive meaning of TransE is to take the relationship r in the triple set (h, r, t) as the translation from entity h to entity t and make $(h + r)$ equal to t as much as possible by adjusting the vectors $h, r,$ and t .

With a given triplet (h, r, t) , TransE model transforms the relation r into a vector r' , so that entities h and r can be connected in the form of vectors with less loss, and the distance function $d(h + r, t)'$ is defined as follows:

$$d(h, r, t) = \|h + r - t\|_2^2. \quad (4)$$

Formula (4) is used to measure the distance between $h + r$ and t , and the hinge loss function is used to minimize it in the training process of the model. The hinge loss function is as follows:

TABLE 1: Training samples based on Skip-gram algorithm.

Input word	Output result 1	Output result 2	Output result 3	Output result 4
Robots	(Robots, walk)	(Robots, around)	—	—
walk	(walk, Robots)	(walk, around)	(walk, boxes)	—
around	(around, walk)	(Around, robots)	(around, boxes)	(around, to)
boxes	(boxes, around)	(boxes, walk)	(boxes, to)	(boxes, the)
to	(to, boxes)	(to, around)	(to, the)	(to, workplace)
the	(the, to)	(the, boxes)	(the, workplace)	—
workplace	(workplace, the)	(workplace, to)	—	—

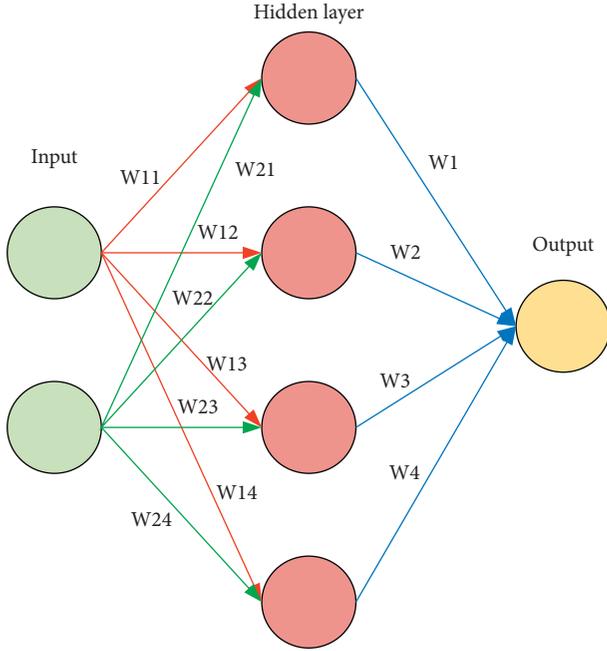
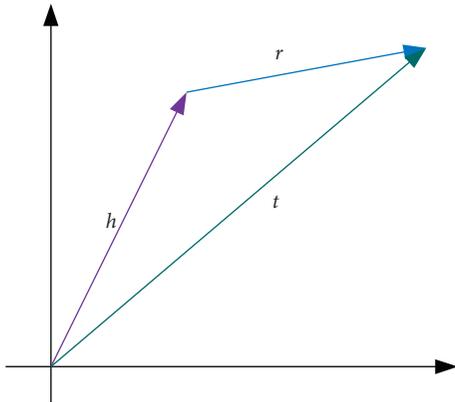


FIGURE 7: Neural network model.



Head + relation ≈ tail

FIGURE 8: Schematic diagram of TransE model.

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'} [\zeta + (h + r, t) - d(\dot{h} + \dot{r}, \dot{t})]_+. \quad (5)$$

In formula (5), s is the triple set in the structured knowledge base, and S' is the triple with negative sampling. ζ is the interval distance parameter, so the value is positive.

$[X]_+$ is a positive function, which is constant when x is a positive number and zero when x is a negative number.

4.4. *Semantic Fusion of Multimodal Entities Based on Singular Value Decomposition (SVD)*. The vector representation of entities based on word2vec and TransE model can solve the semantic problem of entities in single-modal data such as text and structuralization, but cannot solve the knowledge alignment problem between structuralization and text mode. To solve this problem, we get the vector representation set of entity description by word2vec coding the description text (such as color, shape, location, and category) of the entity in structured data as shown in Table 2 and get the text vector description of the entity by averaging the description vectors, so as to realize the entity alignment between structured data and text data, for example, the mobile phone “Apple” is distinguished from the fruit “Apple” by its description information “Block,” “Call,” and “Screen,” and its flow is shown in Figure 10.

After obtaining the multimodal vector representation, the vector dimension is large, which is not conducive to the subsequent similarity calculation. Therefore, it is necessary to reduce the dimension of vector features for calculation.

In this paper, singular value decomposition (SVD) is used to reduce the dimension of features. Input a matrix N can be decomposed into three matrices by SVD, as shown in the following formula:

$$N = A \Sigma B^T. \quad (6)$$

In formula (6), A and B are unitary matrices, Σ is a diagonal matrix, and the singular value of N on the diagonal decreases in descending order. By taking the first m columns and the first m singular values, we can get a new M -dimensional representation, thus realizing dimension reduction.

For multimodal vector representation, the similarity is calculated by calculating cosine similarity, and the calculation formula is as follows:

$$\text{similarity} = \cos(\theta) = \frac{X \cdot Y}{\|X\| \|Y\|} = \frac{\sum_{i=1}^n X_i \times Y_i}{\sqrt{\sum_{i=1}^n (X_i)^2} \sqrt{\sum_{i=1}^n (Y_i)^2}} \quad (7)$$

X_i and Y_i in the formula are the components of vectors X and Y . The similarity of vectors calculated by the above

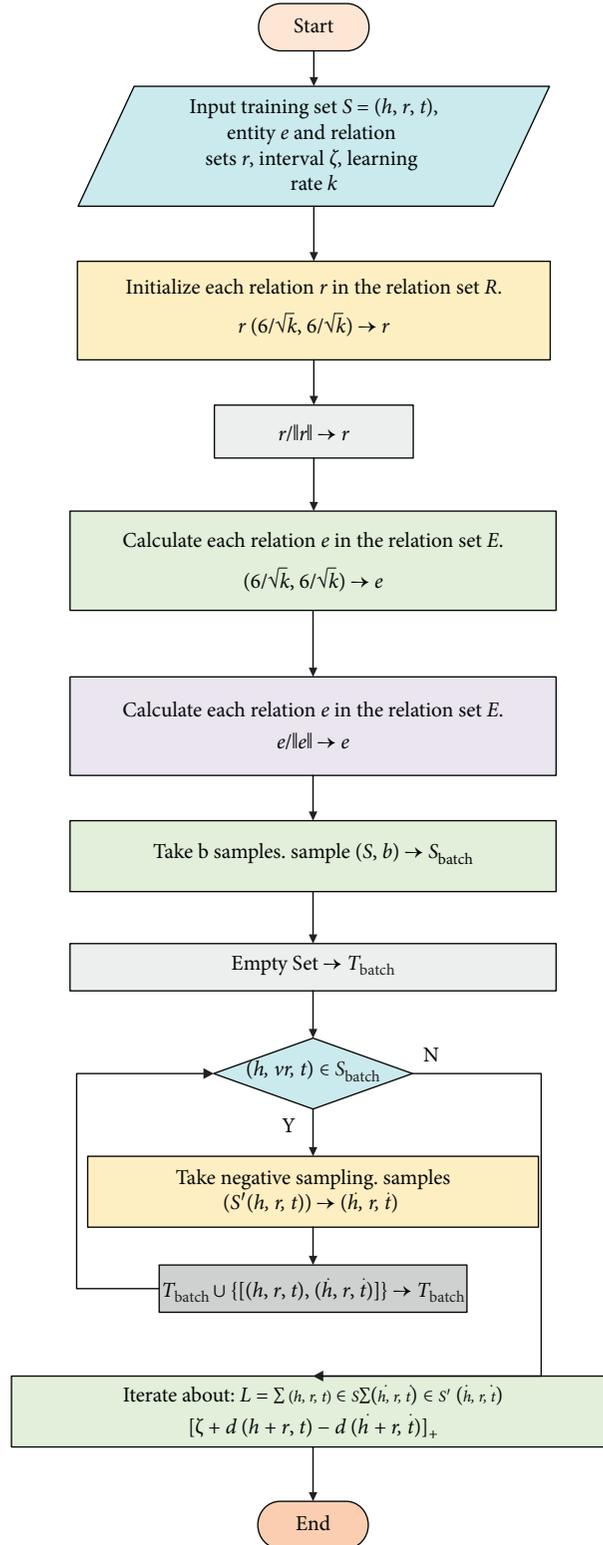


FIGURE 9: TransE algorithm.

formula ranges from -1 to 1 . Among them, -1 means that the entities represented by the two vectors are completely different and 1 means that they are two identical entities. The similarity is between -1 and 1 , and its value represents the semantic similarity between the two entities.

4.5. Knowledge System Fusion Framework Based on HowNet. Semantic alignment of different names of the same entity can be solved by entity semantic fusion algorithm. However, because the knowledge description systems of different knowledge bases are different, some relational knowledge

TABLE 2: Description information of the structural database.

Entity semantic	Description information					
	Color	Shape	Weight (g)	Position	Category
Cucumber	Green	Strip	110	Vegetable warehouse	Vegetable
Apple	Red, green	Oblate	320	Fruit warehouse	Fruit
Orange	Yellow	Oblate	160	Packaging room	Fruit
Workbench	—	<i>n</i> -shaped	5	Packaging room	Equipment

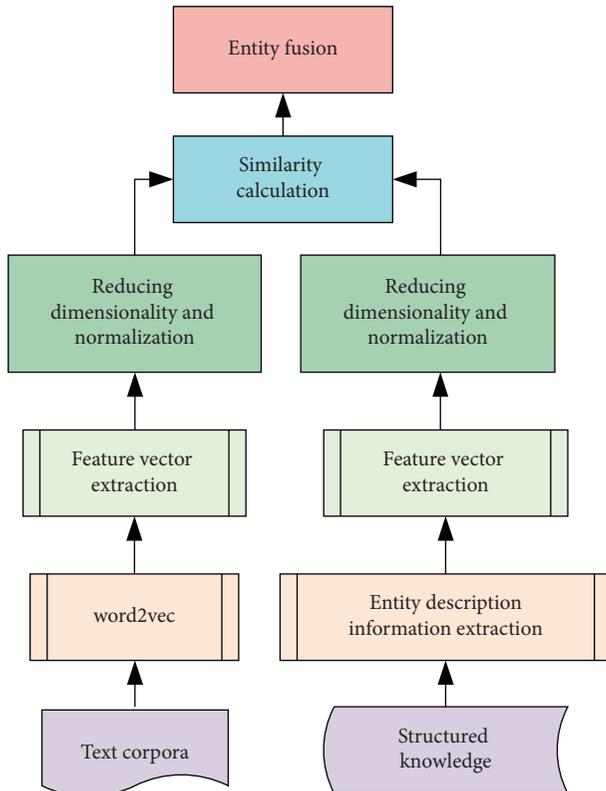


FIGURE 10: Flowchart of multimodal semantic fusion.

bases can only provide the relationship information between entities, while some text knowledge bases focus on the description of entity attributes. In view of the differences of knowledge systems in the above knowledge bases, a multisource semantic knowledge system fusion method is proposed based on Baidu Encyclopedia, Wikipedia, and HowNet semantic dictionary. By providing a unified hierarchical framework of “entity-category-attribute-attribute content” and “entity-relationship-entity,” entity-attribute template and entity-relationship template are established, which effectively solves the problem of multisemantic knowledge base fusion. According to the construction mechanism of knowledge graph mentioned above, the structure of the cognition layer and perception layer is easy to expand data, which makes the robot have the ability to build knowledge base automatically. Knowledge in the perception layer is stored in triples, so that entities, attributes, and relationships can better understand the semantic scope, improve the accuracy of robot item search, and enable robots to have the ability of related search. The flowchart of

knowledge fusion is shown in Figure 11, and its steps are as follows:

Step 1: import the entity set obtained by the knowledge extraction module into the knowledge graph perception layer, fuse the HowNet semantic dictionary, extract the semantic attribute information corresponding to the entities in the semantic dictionary, and then supplement and mark the fused entity attributes

Step 2: learn concept information such as categories from domain documents in encyclopedic knowledge base, use linguistic model to obtain the upper and lower position relationship of categories, use HowNet semantic dictionary to obtain attribute information of categories, and establish cognition layers of categories-attributes and categories-subclasses

Step 3: import the relation set to form a unified normative environmental knowledge representation specification based on the hierarchical framework of “entity-category- attribute-attribute content” and “entity-relationship-entity”

5. Experiments: Results and Discussion

5.1. *Multimodal Entity Semantic Fusion Experiment.* Based on the entity semantic alignment algorithm proposed in this paper, 50 groups of objects related to the robot working environment are extracted from YAGO and encyclopedia knowledge base, respectively. According to the algorithm in this paper, the entity text in encyclopedia knowledge base and the description information in the YAGO knowledge base are represented in the form of vector features. The evaluation metric of similarity calculation method of feature vectors uses the precision rate (p), recall rate (r), and $F1$ values in entity semantic fusion, and their definitions are as follows:

$$P = \frac{TP + TN}{TP + TN + FP + FN},$$

$$R = \frac{TP}{TP + FN},$$

$$F1 = \frac{2 \cdot RP}{R + P}. \quad (8)$$

The meanings of TP, TN, FP, and FN are shown in Table 3.

TP, FP, FN, and TN can be understood as follows:

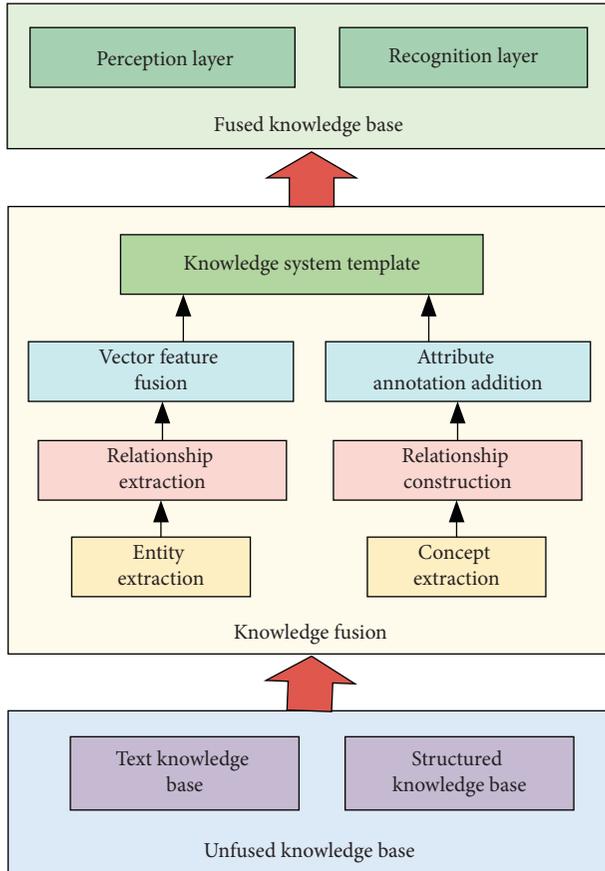


FIGURE 11: Flowchart of knowledge fusion.

- (i) TP: fusion semantic is 1 and actual semantic is 1, fusion correct
- (ii) FP: fusion semantic is 1 and actual semantic is 0, fusion error
- (iii) FN: fusion semantic is 0 and actual semantic is 1, fusion error
- (iv) TN: fusion semantic is 0 and actual semantic is 0, fusion correct

To verify the effectiveness of the algorithm in this paper, we selected the data of eight common fruits and vegetables on the fruit and vegetable sorting line for testing, set the confidence threshold to 0.9, and obtained the confusion matrix of entity semantic fusion accuracy as shown in Figure 12. It can be seen from the confusion matrix that, except tangerines and oranges, whose description information is too similar, which leads to lower accuracy, the accuracy of other semantic fusion has reached a higher level, and the overall performance is good. To further verify the algorithm in this paper, 50 sets of entity semantics and corresponding description information are selected to expand the test data. According to the above evaluation metric of accuracy and $F1$ value, the test results obtained by experiments are shown in Figure 13, with blue broken line as the accuracy result and orange broken line as the $F1$ value result. By observing the semantic fusion results of 50 groups of test samples, the accuracy of the semantic fusion method adopted in this chapter

and the average value of $F1$ value reach about 70%, and the semantic fusion effect is ideal.

5.2. Comparative Experiment on Dimension Reduction of Entity Semantic Feature Vector. At the same time, in order to improve the efficiency of entity semantic similarity calculation, we choose the singular value decomposition method to reduce the dimension of vector and compare it with the vector without dimension reduction in running time and accuracy. The test results obtained by experiments are shown in Figure 14.

By comparing the accuracy curves in Figure 14, we can know the accuracy of entity semantic fusion without dimension reduction and with dimension reduction. Although the accuracy of vector with dimension reduction is slightly lower than that without dimension reduction, they are roughly at the same level. By comparing the running time curves, it can be seen that although the running time of some items in the dimension-reduced vector semantics exceeds that of the no-dimension-reduced vector semantics, the fusion time of the dimension-reduced semantics is much lower than that of the no-dimension-reduced vector semantics. Generally speaking, through the dimension reduction operation of vectors, the matching time is greatly reduced while the success rate of vector matching is kept, and the efficiency of entity semantic matching is improved.

5.3. Cognition Experiment. The main purpose of this experiment is to verify whether the intelligent decision-making ability of the mechanical arm in processing nonspecific complex tasks can be improved after integrating professional database information, functional information, and real-time status information of the mechanical arm into the industrial Internet when constructing the knowledge graph. In the experiment, we placed different fruits and vegetables in three areas around the arm.

As shown in Figure 15, area 1 is an operable area where the mechanical arm can accurately identify the type of items and grab them flexibly. Area 2 is the critical area around the farthest position where the arm can reach the object. Area 3 is the area where the arm must not be able to catch objects. One watermelon and one apple are placed in each of the three areas. We chose a mechanical arm with a grip of 3 kg for the test. The mechanical arm determines the distance of the object by using a depth camera. The instruction we gave was just "take all the fruit and vegetables". During the experiment, each time the arm made a judgment for a grab, we recorded whether it was successful or not and then changed the locations of the watermelon and apple in three areas. A total of 50 experiments were conducted, and the experimental results are listed in Table 4.

In Table 4, the success rate of the mechanical arm in judging apples and watermelons grab in different areas is, respectively, calculated. We found that no matter where the mechanical arm recognized the watermelon, it did not grab it. It shows that the arm knows that it cannot catch watermelon through its knowledge graph. Only in area 3, there were two errors in judgment because the distance was too far and the

TABLE 3: Meaning of TP, TN, FP, and FN.

		Fusion semantic	
		1	0
Actual semantic	1	TP	FP
	0	FN	TN

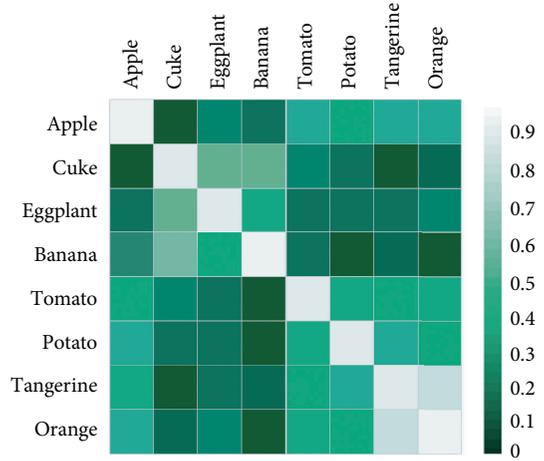


FIGURE 12: Confusion matrix of entity semantic fusion.

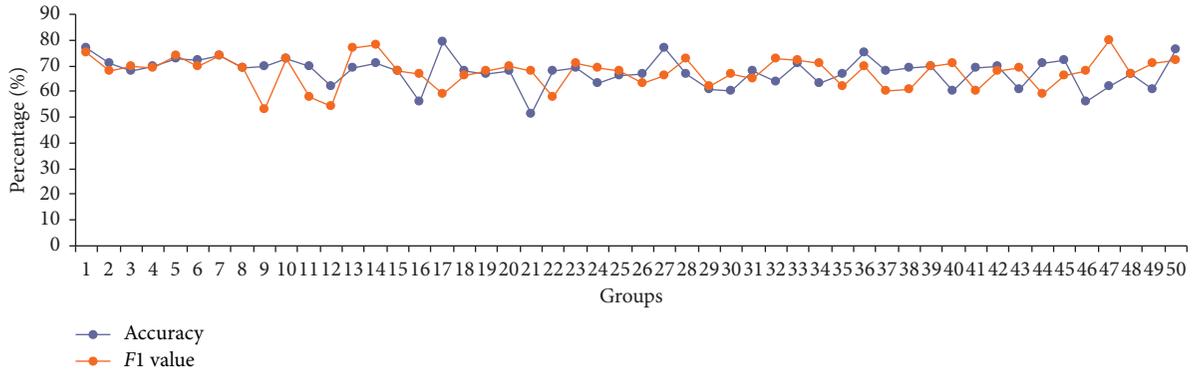
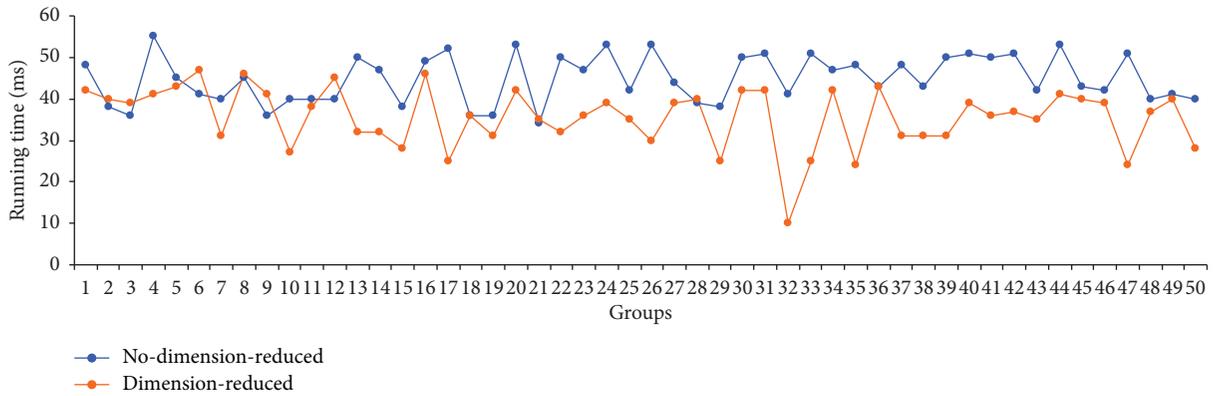
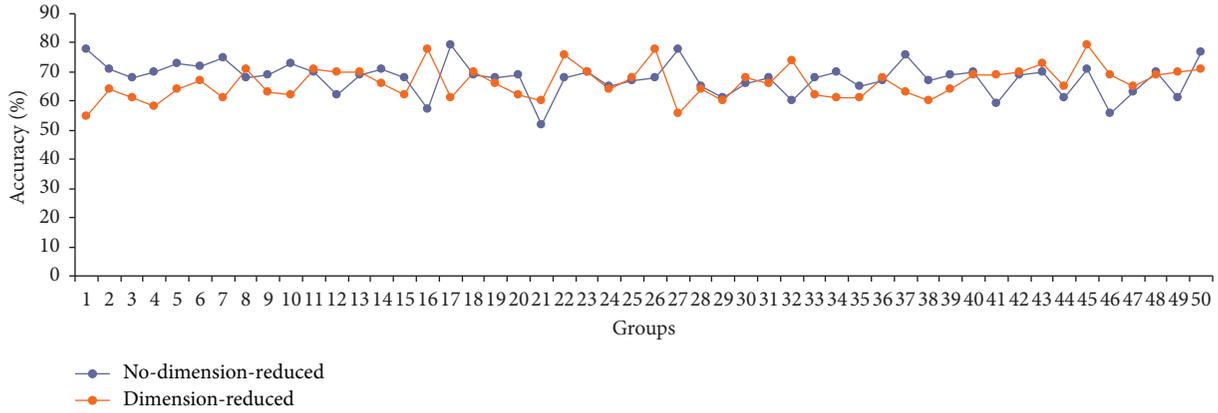


FIGURE 13: Line chart of accuracy and F1 value.



(a)

FIGURE 14: Continued.



(b)

FIGURE 14: Comparison curve of time and accuracy of entity semantic fusion without dimension reduction and dimension reduction. (a) Comparison of running time and (b) comparison of accuracy.

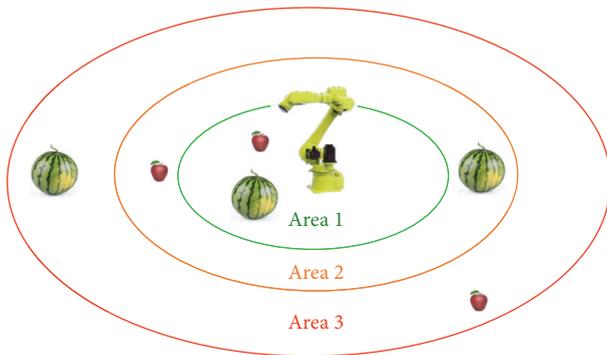


FIGURE 15: Grab the object area diagram.

TABLE 4: Success rate statistics of grab.

Area	Apple (%)	Watermelon (%)
Area 1	98	100
Area 2	68	100
Area 3	90	96

object recognition was wrong. The success rate of recognizing and grabbing apples is the highest in region 1. Only once after successful recognition, grab fails, and the apple slipped. The success rate of recognizing and grabbing apples in region 2 is only 68%. The reason is that when apples are in critical position, the distance judgment accuracy of the mechanical arm is limited, and there are often errors in position recognition. In area 3, the success rate of apple recognition and grab was also high, only 5 apples failed to grab because all the apples in the distance were judged correctly and did not try to grab, while there were 5 apples in area 3 placed close to the critical area, and the arm misjudged that the apples could be caught, resulting in the failure of grab.

6. Conclusions and Future Work

This paper mainly studies how to construct the representation model of environmental information in industrial Internet, proposes a representation model of cognitive information based on knowledge graph, gives the construction process of the model, and describes the detailed construction of the perception layer and cognition layer. After the environmental information is represented by the ontology awareness information and network resource information of the industrial robot and stored in the form of knowledge base, the industrial robot can plan the task according to the position, function, and operation of the items in the task and then perform intelligent operation accurately and efficiently. In addition, we also analyzed the necessity of knowledge fusion and described the problems in data fusion. A concrete fusion scheme is given, which includes entity semantic fusion and knowledge system framework building. The semantic fusion of multimodal entities based on vector representation is realized. It provides a unified hierarchical framework of “entity-category-attribute-attribute content” and “entity-relationship-entity,” which effectively solves the problem of inconsistent knowledge systems in the fusion of multisemantic knowledge base.

In addition, after integrating professional database information, function information, and real-time status information of the mechanical arm into the industrial Internet when constructing the knowledge graph, the ability of the mechanical arm to handle nonspecific complex tasks and the level of intelligence are improved.

This paper has made some achievements in environmental information representation and knowledge fusion, but in the process of knowledge fusion in knowledge base, only the semantic fusion of items is considered, and there is no research on the fusion of items’ relationships and attributes. Therefore, some algorithms related to the fusion of items’ attributes and relationships can be considered to achieve a higher level of knowledge fusion in the future.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Key R & D Program of Shaanxi Province (no. 2019ZDLGY03-04), Xi'an Science and Technology Plan Project (Program nos. GXYD19.8 and GXYD19.7), and the National Natural Science Foundation of China (no. 61303225), and the Space Science and Technology Fund (Program no. D5120200106).

References

- [1] T. Kollar, V. Perera, D. Nardi, and M. Veloso, "Learning environmental knowledge from task-based human-robot dialog," in *2013 IEEE International Conference on Robotics and Automation*, pp. 4304–4309, Karlsruhe, Germany, May 2013.
- [2] S. Lemaignan, R. Ros, R. Alami, and M. Beetz, "What are you talking about? grounding dialogue in a perspective-aware robotic architecture," in *2011 RO-MAN*, pp. 107–112, Atlanta, GA, USA, July 2011.
- [3] S. Lemaignan, R. Ros, E. A. Sisbot, R. Alami, and M. Beetz, "Grounding the interaction: anchoring situated discourse in everyday human-robot interaction," *International Journal of Social Robotics*, vol. 4, no. 2, pp. 181–199, 2012.
- [4] S. Schiffer, N. Hoppe, and G. Lakemeyer, "Natural language for an interactive service robot in domestic domains," *Communications in Computer & Information Science*, vol. 358, pp. 39–53, 2012.
- [5] C. Burghart, R. Mikut, R. Stiefelbogen et al., "A cognitive architecture for a humanoid robot: a first approach," in *5th IEEE-RAS International Conference on Humanoid Robots, 2005*, pp. 357–362, Tsukuba, Japan, December 2005.
- [6] O. Mozos, Z.-C. Marton, and M. Beetz, "Furniture models learned from the WWW," *IEEE Robotics & Automation Magazine*, vol. 18, no. 2, pp. 22–32, 2011.
- [7] D. W. Ko, C. Yi, and I. H. Suh, "Semantic mapping and navigation with visual planar landmarks," in *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 255–258, Daejeon, South Korea, November 2012.
- [8] C. Song, *Research on Building Intelligent Semantic Map of Indoor Robot Based on Distributed Representation*, Beijing University of Chemical Technology, Beijing, China, 2015.
- [9] L. Tamas and L. Cosmin Goron, "3D semantic interpretation for robot perception inside office environments," *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 76–87, 2014.
- [10] H. Wu, G. Tian, X. Chen, T. Zhang, and F. Zhou, "Map building of indoor unknown environment based on robot service mission direction," *Robot*, vol. 32, no. 2, pp. 196–203, 2010.
- [11] H. Wu, "Characterization of large-scale unknown environmental information based on RFID technology," *Journal of Central South University (Science and Technology)*, no. S1, pp. 166–170, 2013.
- [12] W. Park, M. Han, W. J. Son, and S. J. Kim, "Design of scene knowledge base system based on domain ontology," in *2017 19th International Conference on Advanced Communication Technology*, pp. 560–562, IEEE, Bongpyeong, South Korea, February 2017.
- [13] Q. Hao, Y. Li, L. M. Wang, and M. Wang, "An ontology-based data organization method," in *2017 Fifth International Conference on Advanced Cloud and Big Data (CBD)*, pp. 135–140, Shanghai, China, August 2017.
- [14] Y. Yang, Z. Wu, and X. Zhu, "Semi-automatic metadata annotation of web of things with knowledge base," in *2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pp. 124–129, Beijing, China, September 2016.
- [15] P. Das, V. Hilaire, and L. Ribas-Xirgo, *An Ontology to Support Collective Intelligence in Decentralised Multi-Robot Systems*, <https://arxiv.org/abs/1806.00367>, 2018.
- [16] P. Jezek and R. Moucek, "Semantic framework for mapping object-oriented model to semantic web languages," *Frontiers in Neuroinformatics*, vol. 9, no. 3, p. 3, 2015.
- [17] Y. Chen, *Research on Indoor Service Robot Navigation in Large-Scale Complex Scenes*, University of Science and Technology of China, Hefei, China, 2017.
- [18] S. Gao, L. Kong, and P. Wu, "Autonomous processing method of Chinese service instruction for indoor intelligent robot," *Robot*, no. 4, pp. 424–434, 2015.
- [19] A. Singhal, *Introducing the Knowledgegraph: Things, Not Strings*, Official google blog, 2012.
- [20] *Sohu Encyclopedia*, 2015, <http://baike.sogou.com/lh6b61b234.htm>.
- [21] *Baidu Encyclopedia*, 2015, <http://baike.baidu.com/view/10972128.htm>.
- [22] K. Marino, R. Salakhutdinov, and A. Gupta, "The more you know: using knowledge graphs for image classification," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20–28, Honolulu, HI, USA, 2017.
- [23] P. Szekeley, C. A. Knoblock, J. Slepicka et al., *Building and Using a Knowledge Graph to Combat Human Trafficking*, pp. 205–221, Springer International Publishing, NY, USA, 2015.
- [24] Y. Zhang and M. Liu, "Visualization analysis of sports industry research in China based on knowledge map," *China Sport Science and Technology*, vol. 52, no. 1, pp. 24–29, 2016.
- [25] Q. He, "Visualization analysis of physical education research field in China based on knowledge map," *Journal of Beijing Sport University*, vol. 39, no. 2, pp. 98–103, 2016.
- [26] Z. Han, B. Li, K. Zhang et al., "Knowledge structure of China's marine economy research: an analysis based on CiteSpace map," *Scientia Geographica Sinica*, vol. 36, no. 5, pp. 643–652, 2016.
- [27] Y.-H. Xiao, K.-Z. Zhang, and W. Wang, *A Knowledge Graph Construction Method in the Field of Books Reading*, China.
- [28] G. Jin, F. Lü, and Z. Xiang, "Enterprise information integration based on knowledge graph and semantic web technology," *Journal of Southeast University (Natural Science Edition)*, vol. 44, no. 2, pp. 250–255, 2014.
- [29] L. Jia, J. Liu, Y. Tong et al., "Construction of TCM knowledge map," *Journal of Medical Informatics*, vol. 36, no. 8, pp. 51–53+59, 2015.
- [30] X. Lu, J. Zeng, M. Zhang, X. Guo, and J. Zhang, "Using knowledge graphs to optimize MOOC instruction," *Distance Education in China*, no. 7, pp. 79–80, 2016.
- [31] A. Kumar, C. Schmidt, and J. Köhler, "A knowledge graph based speech interface for question answering systems," *Speech Communication*, vol. 92, pp. 1–12, 2017.

- [32] Y. Jia, Y. Qi, H. Shang, R. Jiang, and A. Li, "A practical approach to constructing a knowledge graph for cybersecurity," *Engineering*, vol. 4, no. 1, pp. 53–60, 2018.
- [33] O. Kem, F. Balbo, A. Zimmermann, and P. Nagellen, "Multi-goal pathfinding in cyber-physical-social environments: multi-layer search over a semantic knowledge graph," *Procedia Computer Science*, vol. 112, pp. 741–750, 2017.
- [34] D. Collarana, M. Galkin, I. Traverso-Ribón, C. Lange, M.-E. Vidal, and S. Auer, "Semantic data integration for knowledge graph construction at query time," in *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, pp. 109–116, San Diego, CA, USA, January 2017.
- [35] W. Hao, J. Menglin, T. Guohui, M. Qing, and G. Liu, "R-KG: a novel method for implementing a robot intelligent service," *Artificial Intelligence (AI)*, vol. 1, no. 1, pp. 117–140, 2020.
- [36] J. McCormac, A. Handa, A. Davison, and S. Leutenegger, "Semantifusion: dense 3D semantic mapping with convolutional neural networks," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, Marina Bay Sands, Singapore, May 2017.
- [37] Z. Wang, "RGB-D scene parsing based on spatial structured inference deep fusion networks," *Chinese Journal of Electronics*, vol. 5, pp. 1253–1258, 2018.
- [38] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in Neural Information Processing Systems*, vol. 26, pp. 2787–2795, 2013.