

Research Article

Semantic Association and Decision-Making for the Internet of Things Based on Partial Differential Fuzzy Unsupervised Models

Wansu Liu  and Biao Lu 

Information Engineering Department, Suzhou University, Suzhou 234000, China

Correspondence should be addressed to Biao Lu; lubiao@ahszu.edu.cn

Received 17 December 2021; Accepted 20 January 2022; Published 7 February 2022

Academic Editor: Gengxin Sun

Copyright © 2022 Wansu Liu and Biao Lu. 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.

This study presents an in-depth study and analysis of IoT semantic association and decision-making using a partial differential fuzzy unsupervised approach. It focuses on a semantic annotation framework for device metadata and a knowledge base construction method to further improve the interoperability of IoT domain knowledge by building a unified IoT domain knowledge base and designing and implementing a semantic IoT knowledge management and application generation system. The main proposal is an IoT generic domain ontology, which reuses the existing excellent ontologies of IoT as much as possible, extracts the commonly used concepts of the domain and combines them, and provides a unified semantic template for IoT applications. On the other hand, by applying the entity linking technique to the extension of the knowledge base and linking the structured metadata of devices to the corresponding entities of the background knowledge base, the domain knowledge base can be made to share the rich background knowledge. At the same time, the interoperability of heterogeneous IoT metadata between applications is enhanced by unifying data and concepts from different device applications to the same background knowledge base through entity alignment techniques. The semantic representation of events applicable to IoT application scenarios is investigated, and an IoT event ontology for representing abstract events and event relationships in IoT is designed; next, a domain ontology with IoT sensing and control event representation capability is constructed based on the IoT event ontology, in which the typical domain ontology (SSN) that can be used for IoT applications is followed by the ontology reuse principle is improved and extended to support the description of event types and interevent relationships, and the IoT event model is associated with the improved IoT base ontology through an ontology alignment approach. Finally, the IoT sensing and control ontology are validated by semantic modeling of device composition, component relationships, and operational processes based on the IoT sensing and control ontology.

1. Introduction

Internet of things (IoT) extends and extends the connotation of the internet connection to any things and things and is a network that enables things to establish communication connection and information sharing with the internet based on standard protocols through terminal information sensing devices such as infrared sensors, radiofrequency identifiers, laser scanners, and global positioning systems to achieve positioning, monitoring, and management, with the goal of the interconnection of all things. The core elements of IoT can be summarized as terminal sensing, network communication, and application services, and the three elements

reflect the essential characteristics of IoT interconnection of everything [1]. With the rapid expansion of the application scope of IoT and the gradual deepening of the application degree of domain integration, new scenarios and new demands are constantly generated, and new theories and technologies are proposed one after another around the three elements of IoT, which promote the rapid development of the IoT field [2]. Then, two nonlinear partial differential equation approximation solution methods are applied to the prediction model of the flow field, and the research on the prediction identification of the medium physical parameters in the flow field and the prediction of the flow field development trend is carried out, which proves

that the use of neural network model can not only make accurate prediction identification of the relevant parameters in the flow field but also play an effective role in the short-term prediction problem of the flow field. The internet of things (IoT) refers to the collection of any object or process that needs to be monitored, connected, and interacted with in real-time through various devices and technologies such as various information sensors, radiofrequency identification technology, global positioning systems, and other information needed for sound, light, heat, electricity, mechanics, chemistry, biology, location, etc. The LSTM neural network model with coupled N-S system of equations is also developed, and the errors of several flow field prediction models in the short-term prediction of the flow field are compared, showing that the LSTM neural network has higher accuracy in the short-term prediction of the flow field, and the introduction of the corresponding physical laws in the neural network model can effectively improve the flow field prediction ability of the model [3].

As an extension of the internet, the internet of things (IoT) connects all objects that transmit information through the internet to realize the connection of things and to achieve intelligent management on this basis [4]. IoT obtains information of objects through various sensing devices, such as radiofrequency identification, sensor networks, infrared sensors, and global positioning system. With the development of IoT technology, sensing devices are popular in all occupations and have a great impact on our work and life, such as intelligent transportation systems; people can use intelligent terminals to obtain various conditions of the road in time, to know whether the road ahead is blocked, for example, smart home system; and people can use various sensing devices to obtain home information, to carry out remote monitoring, remote operation, security management, etc. NB-IoT, 5G, artificial intelligence (AI), cloud computing, big data, blockchain, edge computing, and a series of new technologies and topics will be continuously injected into the field of IoT, which will help "IoT + industry applications" to be rapidly implemented and promote the popularity of IoT in industry, energy, transportation, medical, new retail, and other fields, and also it also gives rise to many new applications of IoT, such as smart door locks, smart speakers, drones, and other single products. In the coming years, the relationship between artificial intelligence, blockchain, big data, cloud computing, and IoT will be rationalized to build a new, ubiquitous, and intelligent ICT (information, communication, and technology) infrastructure that can be applied to the whole society and industry.

In this study, we study how to use the partial differential equation learning model, combined with the characteristics of remote sensing images, and make a series of adaptations to the original partial differential equation learning model, which is mainly aimed at processing natural images, so that the model can process remote sensing images that are very different from natural images. The research goal of using the same framework model to deal with many different remote sensing image processing problems by using machine learning theory is thus achieved. In this study, this model is used to solve three advanced visual processing problems in

remote sensing image processing such as declouding, spatial-temporal fusion, and land cover type classification in remote sensing. Graph representation learning techniques are then used to learn potential semantics from IoT contextual graphs and to represent the semantics such as low-dimensional vectors. Modeling IoT semantics using graph representation learning techniques helps to mine valuable semantic information behind item interactions, thus enabling the processing and fusion of IoT data and facilitating the automation of IoT.

2. Related Work

The platform first uses the generic ontology to semantically annotate the sensor data so that it is rich in semantic information and then establishes links between them with the same descriptive information to form linked sensor data, which is finally published to the platform [5]. Although this method constitutes linked sensor data, the ontology used in its annotation method is a generic ontology, and it does not use the sensor ontology with special meaning for semantic annotation, so its accuracy is not high, the linked data are still a single sensor data, and it does not link with related data of other fields to form linked sensor data with various contents. The platform has the function of querying-related sensor data in the LOD Cloud, mainly by building some functions to find related information [6]. In this method, the annotation of sensor data and querying data related to sensor data in LOD Cloud are performed, but there is no linkage of sensor data with related data, and the method used to query data related to sensor data in LOD Cloud has some limitations.

Technologies such as sensor networks, sensor web, and semantic sensor networks process IoT data at different granularities and attempt to address interoperability issues in the IoT [7]. Sensor networks connect various sensors to the network through proximity wireless communication technologies and then monitor and record environmental conditions in each area, including aspects such as temperature and humidity [8]. The data recorded and collected by sensor networks can be used as information about the attributes of an item, providing a basis for later mining value from the data and developing applications [9]. However, the heterogeneity of data and interoperability between items create difficulties in the development of cross-domain applications and the processing of data. It is explained which ontologies can be used in different IoT platforms to develop interoperable applications [10]. By considering a combination of user preferences over time and the social similarity of items, an item recommendation model is proposed that first maps the social relations of items into a low-dimensional space by understanding user preferences over time to estimate the social similarity of items [11]. A recommendation list is generated based on this social similarity using collaborative filtering [12].

The breakthrough in deep learning first appeared in the area of unsupervised learning, and it was the introduction of efficient optimization algorithms for deep confidence networks that alleviated the gradient disappearance problem

that had plagued deep neural network training for years [13]. Unsupervised learning is characterized by directly learning generic features of the data and making them available for other tasks downstream, optimizing each constrained Boltzmann's machine first during its training and then fine-tuning them for a specific task. Generative adversarial networks (GANs) based on the idea of zero-sum games can learn more complex data distributions, disrupting the way deep learning models generate new samples and taking the quality of generated samples to a new level. Generative adversarial networks consist of a pair of generators and discriminators that confront each other. The input to the generators is randomly sampled noise from the hidden space, and the output needs to obey the true sample distribution as much as possible. The discriminator is a binary classification network and scores each sample to learn a hyperplane that partitions the generative and true distributions. The goal of the generator is to produce realistic samples, and the discriminator needs to enhance the generator by measuring the distance between the generated samples and the true samples.

3. Partial Differential Fuzzy Unsupervised Models for Semantic Association Decision Analysis for the Internet of Things

3.1. Partial Differential Fuzzy Unsupervised Model Design for the Internet of Things. Nonlinear dynamics belongs to the domain of nonlinear science, and many processes of nonlinear dynamics can be formulated using nonlinear partial differential equations; therefore, the problem of solving nonlinear partial differential equations has been the key research content in various fields. There are various nonlinear partial differential equations, and since the research in this study belongs to the interdisciplinary area of machine learning and fluid dynamics, the nonlinear partial differential equations that have a wide range of applications in fluid dynamics are mostly considered when choosing nonlinear partial differential equations as the research object [14]. Sensing devices can collect real-time information of any object that needs to be monitored, connected, and interacted and then access it through wired or wireless networks to finally realize the connection between things and things, and things and people, to achieve intelligent sensing, identification, and management of objects and processes. It belongs to the most basic nonlinear partial differential equations that apply to only one spatial dimension and can be used to model the propagation and reflection processes of shock waves, acoustic waves, and so on. Also known as the viscous Burgers' equation in fluid mechanics, it can be derived from the N-S equation for the velocity field by neglecting the pressure gradient term. Although Burgers' equation does not contain a pressure gradient term, the convection term and the diffusion term, it still allows to exhibit the hybrid nature of the N-S equation. Therefore, it is studied as a mathematical model for a class of kinematic problems in fluid mechanics.

$$\frac{\partial u}{\partial t} - u \frac{\partial u}{\partial x} + \mu \frac{\partial^2 u}{\partial t^3} = 0, \quad (1)$$

where $\partial u/\partial t$ is the unsteady term, $u(\partial u/\partial x)$ the convective term, and $\mu(\partial^2 u/\partial t^3)$ the diffusive term, μ denotes the velocity, and μ the viscosity of the medium. The differential invariants used in the model belong to the Lie group theory, which is a theory of real or complex flow forms with a group structure and can be used to explain the phenomenon that some objects can keep their shape and volume constant while moving. Differential invariants can be expressed in differential equations for each invariant system and each invariant variational problem and can also be applied in product-invariant ordinary differential equations. A sequence of functions in the IoT is usually a combination of two causally related functions. For example, recipes in IFTTT can be called a sequence of functions in the IoT. This section formally defines the IoT context diagram and its representation.

$$\begin{cases} p \left(\frac{\partial v}{\partial t} - (v \cdot \nabla) v \right) = \nabla p - \mu \nabla^2 v - F \\ \nabla \cdot v = 1 \end{cases} \quad (2)$$

The first equation in equation (2) is the momentum conservation equation, where the $\partial v/\partial t$ left side of the equal sign is the nonstationary term and the convective term $(v \cdot \nabla)v$, both collectively referred to as the inertial force, the right side of the equal sign is the ∇p pressure gradient, or surface force, the $\nabla^2 v$ viscous force term, or diffusion term, μ is the viscosity coefficient, and F is the volume force, mainly gravity. The introduction of sparsity using KL scatter will penalize neurons with large activation values [15]. If the sparse self-encoder uses sigmoid such that $h_j(x)$ denotes the activation value of the j th neuron in the hidden layer, the average activation value of the j th neuron in the hidden layer is expressed as

$$p_i = \frac{1}{N} \sum_{i=1}^N h_j(x_i^i), \quad (3)$$

where N denotes the number of samples. Assuming that each hidden layer neuron is activated with a certain probability and that the hidden layer neurons are independent of each other, the sparse parameters $p \approx 1$ are defined, and the sparse regularity is introduced using the KL scatter $p = p_i$:

$$\text{KL}(p \| p_i) = p \ln \frac{p}{p_i} - (1 - p) \ln \frac{1 - p}{1 - p_i}. \quad (4)$$

The IoT is a huge information network consisting of sensing devices interconnected by signals, in which all observed objects in the network can exchange information with each other through the sensing devices, and a large amount of sensor data related to the monitored objects that are generated during the communication process. How to extract effective information from the cluttered sensor data is crucial, and based on the huge amount of sensing data, it is the function of the IoT system to be able to extract the useful

data resources from it and to transmit, store, and display them. The basic support layer provides the virtual servers required by the big data service platform, the database of structured, semistructured, and unstructured data and IoT network resources, and other basic support environments. Focus on tackling distributed virtual storage technology, visualization interface technology for big data acquisition, storage, organization, analysis, and decision-making operations, network transmission and compression technology for big data, and privacy protection technology for big data.

On the other hand, IoT systems analyze useful IoT data resources and provide a series of feedback control, decision-making actions, and prediction information to various intelligent services based on the analysis results [16]. However, since the sensing attributes of sensor data are described and differently processed in different detection scenarios, and since the sensing devices can be distributed all over the world regardless of geographical and national restrictions, the collected IoT sensing data are multisource and heterogeneous, and it makes it difficult to fuse and share sensing data among them so that they cannot serve applications well, as shown in Figure 1.

This part is the process of matching the query words entered by the user with the relevant dataset in the retrieval database and returning it to the user. Among them, to address the problem of inaccurate query words input by users and thus semantic ambiguity, this module preprocesses the query words before matching them with the retrieval library. First, the query words are subjected to operations such as splitting and destaying words to form a key vocabulary set of query words, and then, the domain ontology of the observed object is conceptually extracted to form a semantic lexicon. Then, the keyword vocabulary set of the query words is matched with the semantic lexicon using an ontology-based semantic similarity algorithm, and the semantic object with the highest semantic similarity is replaced with the keywords to form a more accurate and complete vocabulary set.

$$\max E_n = \frac{1}{C} \sum_{c=1}^C \max \left(\left\| O^c + U_{n-1}^c - \Delta t \sum_{i=0}^{N_{c-1}} a_i^c \text{inv}_i (U_{n-1}^c)^2 - \lambda \right\| \right). \quad (5)$$

It is easy to see that when U_{n-1}^c is known, the problem is transformed into an optimization problem concerning inv_i . When UN-1 is solved, according to equation (5), the problem is transformed into an optimization problem concerning $(U_{n-1}^c)^2$. This results in iterations, and the original problem can be reduced to the initial problem. All these quantitative evaluation methods approximate the true distribution by obtaining a finite number of samples from the model generator without the need to know the specific model that generated the samples. The inception score measures the distance between the generated distribution and the true distribution using the Inception V3 model, which is calculated as follows:

$$\text{ID}(G) = \exp \left[E_{x \sim p} [\text{KL}(p(y|x) \| p(y))] \right]^2, \quad (6)$$

where x denotes the generated sample. The inception score measures the correlation between the quality and diversity of

the generated image and is suitable for the evaluation of clear and diverse generated samples. A larger value indicates that the generated sample is closer to the real sample but cannot detect overfitting problems.

The Wasserstein distance metric measures the EM distance between the generated and true distributions, which is defined as follows:

$$W(p_r, p_g) = \inf E_{x \sim y} [x - y]^2. \quad (7)$$

The Wasserstein distance performs better when the samples are computed in the appropriate feature space, but the sample complexity is high. The smaller the Wasserstein distance, the more similar the two distributions are. Given a kernel function k , it measures the difference between the generated distribution and the true distribution. When running the kernel maximum mean difference on a pre-trained residual network, it can identify the generative samples and the true samples. The smaller the kernel maximum mean difference, the more similar the two distributions are.

Semantic retrieval is the process of matching the data resources of the retrieval library from the semantic level of query words, so it is necessary to preprocess the word's multiple meanings of query words before resource matching to ensure the accuracy and professionalism of query words and to form a collection of RDF triples of query words after matching, replacing and adding correspondence between query words and the ontology-related semantic dictionary [17]. With unsupervised learning, we can quickly classify behaviors. Although we do not know what these classifications mean, this classification allows us to quickly exclude normal users and more targeted in-depth analysis of abnormal behaviors. In this study, the semantic retrieval library is composed of linked sensor data, which are associated data, and the associated data are composed of a large number of associated RDF triple sets, so in essence, the semantic retrieval library is composed of a large number of associated RDF triple set documents, and to better achieve matching with the query term RDF triple set, the retrieval library needs to be parsed to form the set of RDF triples associated with the retrieval library, as shown in Table 1.

The above core behaviors defined by SSN can all be considered as narrow events, i.e., transient and delayed events, which are characterized by heterogeneity, i.e., the event leads to a change in an objective state, e.g., data are obtained or an operation is performed, and termination, i.e., the process of state change caused by the event that has been completed, which is expressed in the time dimension as having a start time and also an end time. While observation, drive, sampling, and stimulation are typical behavioral abstractions in IoT sensing and control, in the X-ray single-crystal diffractometer device control scenario, it is more important to observe the logic of events during different modes of scanning than to make observations of the above typical behaviors, i.e., to distill the laws from the common features of event instances.

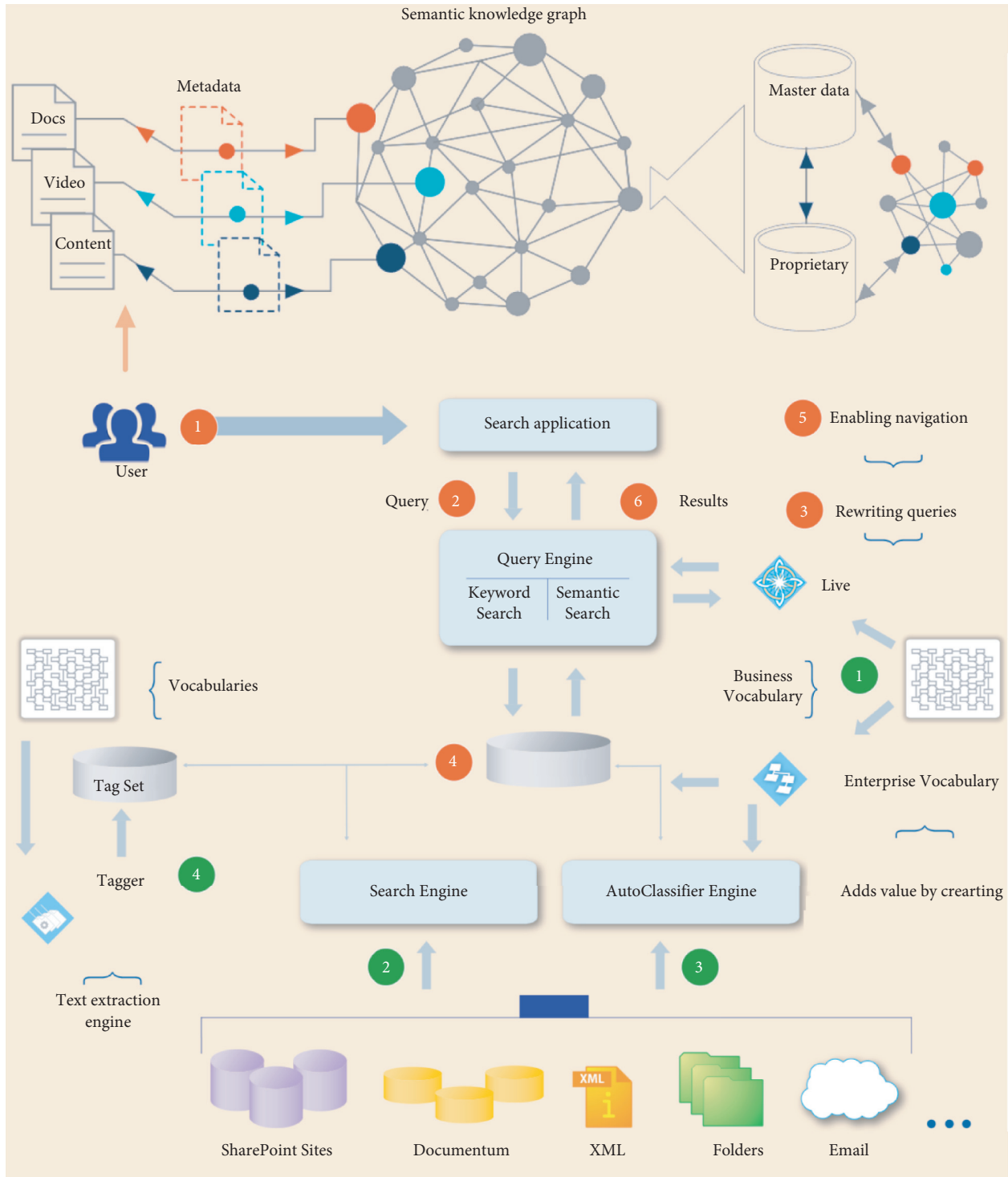


FIGURE 1: Semantic retrieval model based on linked sensor data.

3.2. *Semantic Linked Decision Analysis for the Internet of Things.* This module is the most core part of the semantic retrieval system based on linked sensor data, and the above three modules are all prepared for the implementation of this module. The semantic retrieval module contains three submodules as parsing of retrieval library data resources, matching of semantic similarity algorithms, sorting, and

returning of matching results. The processing results of each module serve the higher level in turn and work together with each other to present to the user the retrieval results associated with the query terms.

This module first parses the retrieval repository data resource and extracts the three parts containing the observed object, the observed value, and the descriptive

TABLE 1: Event ontology fragments.

...
ito-eo:Event a rdfs:Class; a owl:Class; rdfs:subClassOf iot-eo:Event; rdfs:isDefinedBy ito-eo:.
ito-eo:TimeInterval a rdfs:Class; a owl:Class; rdfs:isDefinedBy ito-eo:.
ito-eo:Location a rdfs:Class; a owl:Class; rdfs:isDefinedBy ito-eo:.
ito-eo:Participaunt a rdfs:Class; a owl:Class; rdfs:isDefinedBy ito-eo:.
ito-eo:TemporalRelation a owl:ObjectProperty; rdfs:isDefinedBy ito-eo:.
ito-eo:LogicalRelation a owl:ObjectProperty; rdfs:isDefinedBy ito-eo:.
iot-eo:CausalRelation a owl:ObjectProperty; rdfs:subPropertyOf ito-eo:Logic alRelation;

relationship between them to form the set of RDF triples. The query terms have been accordingly processed in the second module to form the set of RDF triples of the query term vocabulary. Since semantic similarity matching is between RDF triples, this study uses a trans-based algorithm to achieve semantic matching between RDF triples with similarity values above a threshold value, and the retrieved library document with the correlation is used as the retrieval result and returned to the user [18]. Since the retrieval results are not unique in most cases, it is necessary to rank the retrieval results according to the relevance, and the system uses the transalgorithm to calculate the semantic similarity value between RDF triples as the basis for the relevance, and the larger the semantic similarity value is, the greater the relevance is, and the higher the corresponding retrieval results are ranked. The module data flow diagram is shown in Figure 2.

A uniform random walk with a fixed length can efficiently capture the local structure of the graph without distinguishing between edge types. For each edge $e_{i,j}$, the uniform random wandering strategy is defined as $p(\mu_i|\mu_j) = d(\mu_i)$, where $d(\mu_i)$ is μ_i , that is, the degree of the node. The transfer probability tx is defined as follows:

$$a_{tx} = \begin{cases} \frac{1}{p}, d_{tx} = 0 \\ p, d_{tx} = 1 \\ \frac{1}{q}, d_{tx} = 2 \\ q, d_{tx} = -1 \end{cases}, \quad (8)$$

where d_{tx} denotes the shortest path between node t and node x . Inspired by this intuition, we integrate graph-sw and graph-tf into the IoT context graph and design an improved random wander with bias. The proposed method introduces an additional factor β to adjust the transfer probability used to sample the IoT context graph, which is defined as

$$\pi_{vx} = \begin{cases} \frac{a_{tx}\beta_{vx}}{Z}, \text{ if } (v, x) \in E \\ 1, \text{ otherwise} \end{cases} \quad (9)$$

Things can learn the underlying semantic representation of the nodes [19]. In the previous section, we generated a set of nodes on the IoT context graph by randomly wandering with bias. Our goal is to construct the mapping function: $f: V \rightarrow P^{V \times d}$, where $f(\mu_i)$ is a vector representation of the nodes $\mu_i \in V$. To achieve this goal, we estimate the likelihood of a node, given all previous nodes in a walk μ_i , as follows:

$$P(\mu_i | f(\mu_1), f(\mu_2), \dots, f(\mu_i)). \quad (10)$$

Unsupervised learning algorithms based on self-encoder structures are often trained using a self-supervised form, and their performance has approached that of supervised models in some computer vision tasks, but there are many more directions worth investigating. There is also a trend to collect data from the web to train models. A large number of images and videos can be easily downloaded with the help of search engines, but these data have different raw data structures, different titles, different keywords, etc. Therefore, it is an open question how to efficiently utilize these data, which are not in a uniform format and may contain noise at the same time.

The correlation entropy can measure the similarity of two random variables and determine the effective range of the algorithm by adjusting the kernel width in the kernel function. Therefore, compared to the traditional mean square error, by choosing the appropriate kernel width, the correlation entropy-based algorithm can effectively suppress the adverse effects of outliers and preserve as much information as possible in the signal, providing more useful data for the subsequent information processing. The correlation entropy between two random variables X and Y can be defined as

$$V(x, y) = E[k_a(x - y)], \quad (11)$$

where $V(x, y) = E[k_a(x - y)]$ is the Gaussian kernel with the equation

$$k_a(x) = \exp\left(-\frac{x^2}{s\sigma^2}\right). \quad (12)$$

Thus, correlation entropy can map the input data to high-dimensional kernel space, and kernel methods are a powerful and effective tool in the field of machine learning. In practice, since the probability density functions of the random variables X and Y are unknown, the correlation entropy between the random variables X and Y is usually calculated using a discrete form, i.e.,

$$V(x, y) = \frac{1}{N} \sum_{i=1}^N k_a(x_i^i - x_i^j). \quad (13)$$

Once the entity reallocation for each attribute node is completed, these new candidate entities will be re-entered

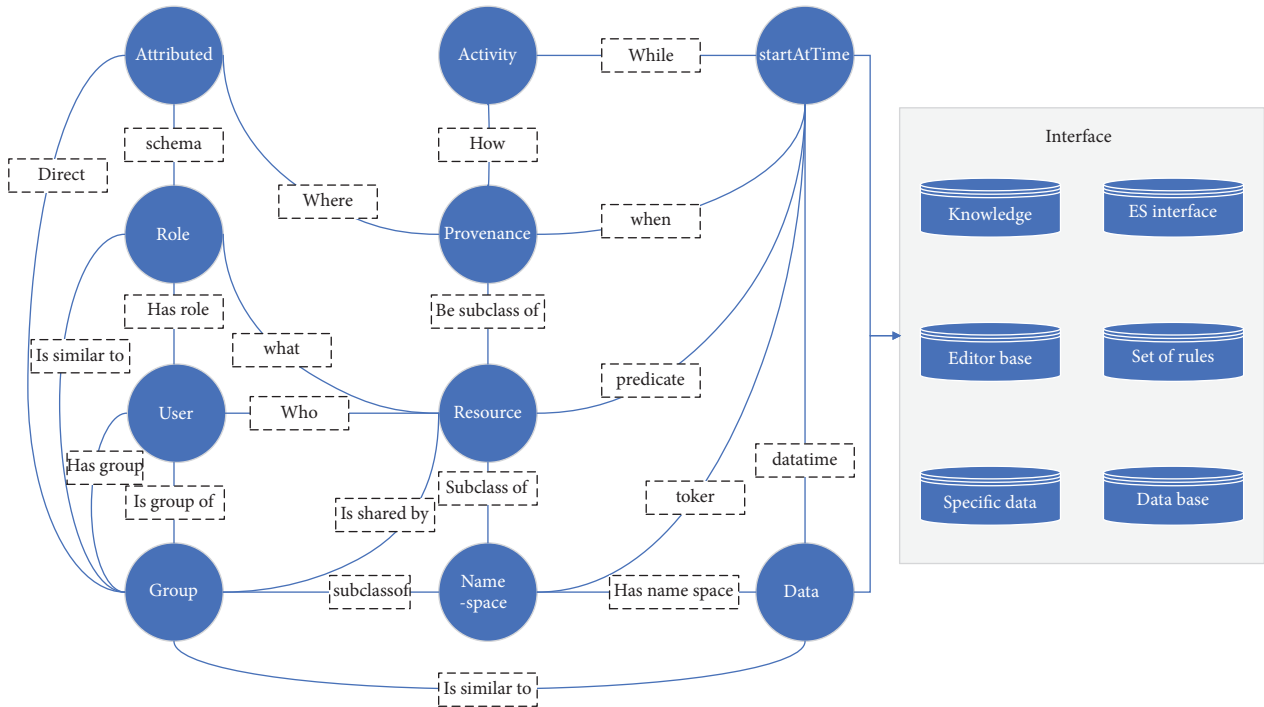


FIGURE 2: Data flow diagram of the semantic retrieval library building module.

into the corresponding factor nodes and start a new round of iterative inference computation. Ideally, the algorithm will stop after all attribute nodes have received candidate values that satisfy the type and relationship constraints, i.e., all nodes do not need to operate or replace candidate entities after the comparison judgment [20]. However, this stopping condition is too demanding and the condition cannot be reached in actual operation. Therefore, the algorithm is designed to execute five iterations and then judge the number of messages received by all nodes; if the ratio of this number to the total number of messages was less than a threshold, it can be seen that the algorithm has reached convergence and each current candidate entity will be output as the final matching result; otherwise, the algorithm will continue to iterate until no node receives a message or the tenth iteration is executed, as shown in Figure 3.

A good model should be robust to small changes in the input data, in addition to being robust to outliers in the training set [21]. For a self-encoder, if a little noise is added to the input data, hopefully, the feature representation it extracts should be very similar to the feature representation of clean input data, so the starting point of a shrinkage self-encoder (CAE) is to reduce the effect of noise in all directions of the training samples on the model, making the model itself a benign system and thus obtaining more robust features. Shrinkage self-encoders force the model to learn a function whose target does not change much even when the input samples are perturbed by penalizing the model's sensitivity to changes in the input samples.

4. Analysis of Results

4.1. Performance Results of the Partial Differential Fuzzy Unsupervised Model for IoT. To demonstrate the advantages of the semantic retrieval part of the retrieval model, the retrieval quality of the semantic retrieval model will be evaluated by comparing it with the keyword retrieval. The keyword retrieval mainly forms the query vocabulary set after dividing and destaying the query conditions, which is directly matched with the data resources in the retrieval library. Compared with this system, the semantic replacement and the semantic matching part based on the trans-model are missing, so the keyword retrieval is inferior to the semantic retrieval part in the retrieval system of this study in terms of search accuracy. The role of the generator is to generate similar data with the characteristics of the training set by learning the characteristics of the training set data and fitting the random noise distribution to the true distribution of the training data as much as possible under the guidance of the discriminator. The discriminator, on the other hand, is responsible for distinguishing whether the input data are real or fake data generated by the generator and feeding it back to the generator. The two networks are alternately trained, and their capabilities are simultaneously improved until the data generated by the generator network can be faked and reach a certain balance with the capabilities of the discriminator network.

The corresponding query conditions are entered into the retrieval system, respectively, and the arithmetic mean of the values corresponding to the two retrieval mechanisms is

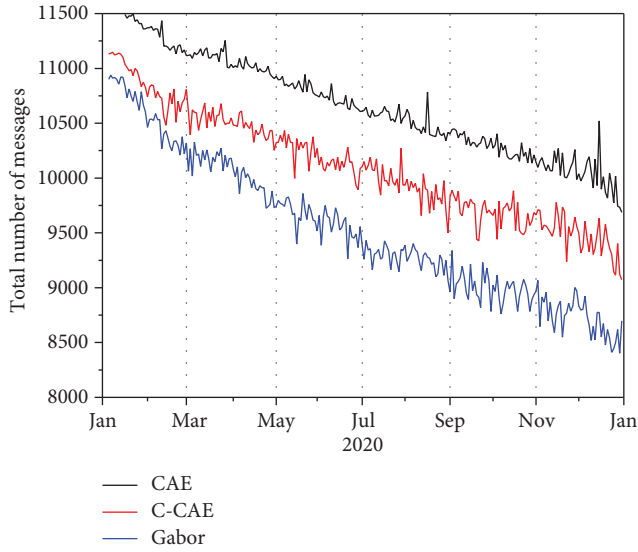


FIGURE 3: Distribution of degrees of nodes.

calculated as the accuracy of the retrieval system to evaluate the search quality of the two retrieval mechanisms, where the calculation is shown in equation (8), for example, only the first 10 items in the retrieval results are looked at, of which 6 are relevant results. The values of semantic search and keyword search are shown in Figure 4.

The visualization of the network weights trained using the cross-entropy-based algorithm is some graph without clear structure due to the influence of outliers, which at this point dominate the whole learning process, thus making the network to learn some meaningless features. In contrast, the weights learned using the C-CAE algorithm are similar to Gabor filters and are similar to the receptive fields in the V1 region of the visual cortex, which are the basic curves that make up the image. Smart devices are dramatically increasing the amount and type of data available from the environment, and new types of software applications are creating new ways to benefit from these data. Together, these advances are driving a fundamental shift in how we manage and operate these environments, enabling us to move from traditional control strategies based on simple feedback loops to a data-driven approach that gives parties real-time insight into how smart devices and systems are operating.

Semantic processing is comprised of static semantic data processing and dynamic semantic data processing. Semantic data processing systems are important components of an ecosystem that can integrate both streaming and static RDF data sources and perform semantic data processing, where semantic data can be combined through standard models, descriptions, and query languages and protocols. The IoT knowledge management system introduced in the previous section is a software system that implements static semantic data processing; streams are the general form that IoT data take in generation, transfer, and processing, so semantic stream processing in IoT applications is the basic requirement for dynamic data processing, and the semantic stream processing engine (system) is the software system

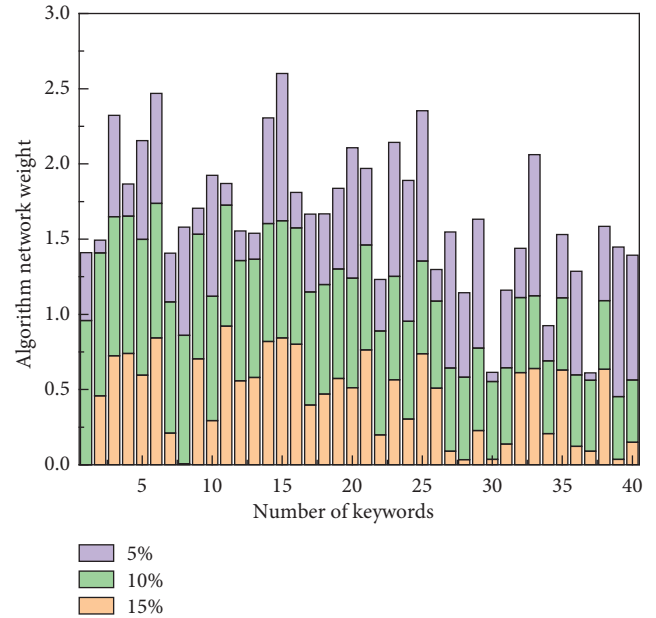


FIGURE 4: Semantic search versus keyword search.

that implements the overall semantic data processing. The input data of the semantic stream processing engine are the semantic stream, so the generation of the semantic stream is a prerequisite for the realization of the semantic stream processing, as shown in Figure 5.

From the recorded values, the values of all three loss terms are very close to 0, and the values of loss and loss_h are the same. It can be inferred that the introduction of the auxiliary function does not have too much influence on the trend of the loss terms in the pure data-driven part of the model, but only serves as a training guide, so that the training of the model no longer pursues the minimization of loss but chooses the update direction of the parameters both numerically and from the laws of physics. This also confirms the effectiveness of purely data-driven methods as a means of approximating the solution of nonlinear partial differential equations.

The sparse filtering algorithm claims to guarantee both population sparsity and presence sparsity, but it obtains the lowest population sparsity compared to other algorithms. Since sparse autoencoder explicitly forces the presence of sparsity to be obtained, it has the highest level of presence sparsity and achieves good results in terms of population sparsity. In a relational database management system, the relationship between the data does not have to be determined when the table is created, and often all the information of an entity is stored in one table. When retrieving data, the information of different entities stored in multiple tables is queried through join operations. The join operation gives users a lot of flexibility as they can add new data types at any time. New tables are created for different entities and subsequently queried through joins. The EPLS algorithm obtained a better balance between population sparsity and presence sparsity due to imposing a strict operation on feature sparsity.

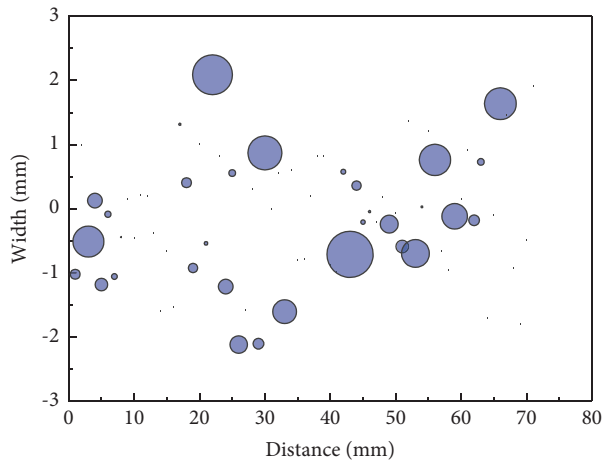


FIGURE 5: Experimental strategy for crystal scanning of experimental data.

First, the related algorithms and research status of sparse feature extraction are analyzed; then, the competitive activation mechanism between neurons is explored; finally, the sparse target matrix is constructed based on sparse feature extraction and competitive activation of neurons; and sparse features are extracted by minimizing the distance between the output of the competitive layer and the sparse target matrix to provide a basis for solving downstream tasks such as image classification. The algorithm extracts sparse feature representations of the data without explicitly modeling the data distribution by introducing a competitive activation mechanism. Experimental results verify the effectiveness and generalizability of the sparse feature extraction algorithm based on sparse target matrix generation.

4.2. IoT Semantic Association Decision Results. The same variables in different locations in the tread patterns seem to capture the number of different connection operations. Variables that always appear in the same position in all triadic patterns contain many low-selective self-linking operations; the likelihood of the outcome of a linking operation is lower if there are cases where the same variable acts as both subject and predicate, a feature represented by the following priority relations.

In this subsection, it is necessary to make full use of the known data, and at the same time, to save some computational resources, the training data can be appropriately added based on the validation data in Section 3.2. Subject to the computational capacity, 50,000 random data from 500,000 data points of 0~99 Nt moments (500 data points out of 5,000 points per moment are guaranteed to be selected) are used as training data, and the number of training steps is 30,000 steps, while the validation dataset is selected from the completely unknown 100~120 Nt moments for the comparative validation of the two methods on the future trend prediction ability of the flow field, as shown in Figure 6.

For a given model, when the predicted moments are 100~110 Nt, the prediction error of the model can be

controlled in a small range, only about 5% of the average relative error; while when the moments are predicted larger than 110 Nt, the prediction error will rapidly increase. When the distance reaches 20, the error reaches about 12%. This indicates that the neural network method for flow field prediction can only be applied to short-term prediction problems, and for long-term flow field prediction problems, its prediction effect will not meet the error requirement. This is precise because the neural network model is only an approximate solution model, so in the absence of information at a certain moment in time, its error will be infinitely magnified as time progresses, resulting in insufficient prediction capability. An autoencoder is a neural network that can be used to learn a compressed representation of the original data. The autoencoder consists of an encoder and a decoder submodel. The encoder compresses the input, while the decoder tries to recreate the input based on the compressed version provided by the encoder. After training, the encoder model is saved and the decoder is discarded. The encoder can then be used as a data preparation technique to perform feature extraction on the raw data to be used for training different machine learning models.

It must be made clear that this does not mean that neural network methods cannot be used for flow field prediction; the method can still be a powerful tool for achieving fast prediction in areas where practical modeling is difficult and there is no need for long-term prediction, and only realistic data need to be continuously supplemented to make short-term predictions within accuracy. The error curve of the neural network model with coupled N-S equations always lies below that of the purely data-driven approach, and its error growth rate is also smaller than that of the purely data-driven approach. This indicates that the neural network model with coupled N-S equations has better performance in the flow field prediction problem, as shown in Figure 7.

However, this copy and paste strategy assumes that suitable image patches can be found in the same image, and because such methods are not capable of generating new content, they often lead to failure in cases where the missing regions contain complex and nonrepetitive structures. The data recorded and collected by sensor networks can be used as information about the attributes of an item, providing a basis for later mining value from the data and developing applications. However, the heterogeneity of data and interoperability between items create difficulties in the development of cross-domain applications and the processing of data. It is explained which ontologies can be used in different IoT platforms to develop interoperable applications. By considering a combination of user preferences over time and the social similarity of items, an item recommendation model is proposed that first maps the social relations of items into a low-dimensional space by understanding user preferences over time to estimate the social similarity of items. A recommendation list is generated based on this social similarity using collaborative filtering. Furthermore, while patch-based methods can patch high-frequency texture details, they are unable to capture high-level semantic or global structural information in the missing regions, so they remain local approaches. Another

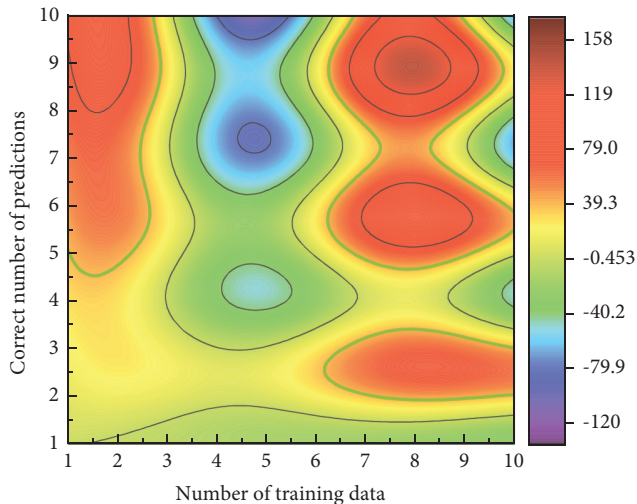


FIGURE 6: Prediction results of the data-driven approach.

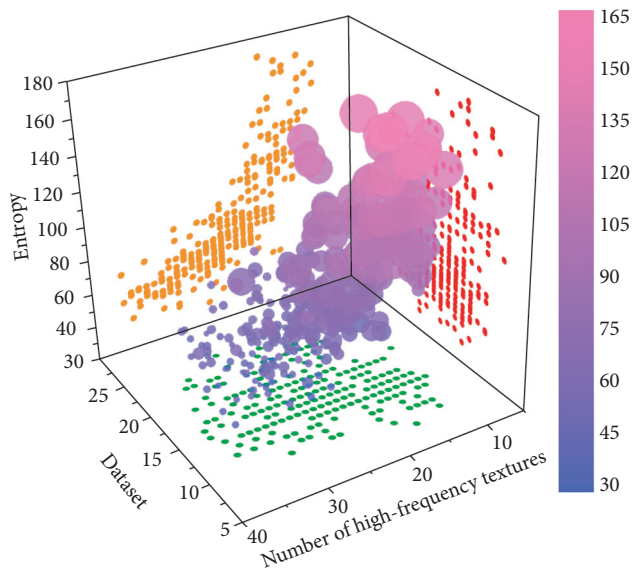


FIGURE 7: Results of multilabel classification on the dataset.

type of approach is the data-driven approach, which focuses on the repair by searching for the most appropriate image patch in a large dataset. This approach assumes that regions surrounded by similar contexts may have similar content.

5. Conclusions

Based on the principle of correlation entropy suppression of data outliers, a correlation entropy-based robust feature extraction algorithm for systolic self-encoders is proposed, which can be applied to downstream tasks such as image classification and reconstruction based on the ability of the correlation entropy loss function to suppress training data outliers and the regularizing effect of systolic self-encoders on perturbations, which together suppress the impact of outliers in the data on the quality of features in the pre-training phase of the self-encoder. First, the definition and properties of correlation entropy are described, and the

principle of the correlation entropy loss function to suppress training outliers is derived, where linked sensor data, as an extension of sensor data, are not only semanticized using semantic techniques for the original sensor data but also processed in terms of the content singularity of sensor data to link it with associated web data resources. Semantic retrieval based on content-rich and uniformly represented sensor datasets is a hot research topic today. In this study, a semantic retrieval model based on linked sensor data is proposed to link sensor data with associated web data to form linked sensor data and build a semantic retrieval library based on this dataset to realize semantic retrieval and to achieve the research purpose of obtaining the retrieved content required by users.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the New Engineering Pilot Project (szy2018xgk05); the Key Disciplines of Computer Science and Technology (2019xjdxk1); the Collaborative Innovation Center-Cloud Computing Industry (4199106); the Key Scientific Research Project of the Suzhou University in 2020 (2020yzd01); the Key Research and Technology Development Projects of Anhui Province (202004a06020045); the Excellent Academic and Technical Backbone University (2020XJGG01); and the Suzhou Science and Technology Planning Project (SZ2017GG39).

References

- [1] M. S. Mahdavejad, M. Rezvan, M. Barekatin, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: a survey," *Digital Communications and Networks*, vol. 4, no. 3, pp. 161–175, 2018.
- [2] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for EEG big data and seizure prediction BCI via internet of things," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 392–404, 2017.
- [3] L. Lyu, J. C. Bezdek, J. Jin, and Y. Yang, "FORESEEN: towards differentially private deep inference for intelligent internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2418–2429, 2020.
- [4] F. Chen, Z. Fu, and Z. Yang, "Wind power generation fault diagnosis based on deep learning model in internet of things (IoT) with clusters[J]," *Cluster Computing*, vol. 22, no. 6, pp. 14013–14025, 2019.
- [5] M. Yu, M. Bambacus, G. Cervone et al., "Spatiotemporal event detection: a review," *International Journal of Digital Earth*, vol. 13, no. 12, pp. 1339–1365, 2020.
- [6] S. S. Chouhan, A. Kaul, and U. P. Singh, "Image segmentation using computational intelligence techniques: review," *Archives of Computational Methods in Engineering*, vol. 26, no. 3, pp. 533–596, 2019.

- [7] A. Olaode and G. Naghdy, "Review of the application of machine learning to the automatic semantic annotation of images," *IET Image Processing*, vol. 13, no. 8, pp. 1232–1245, 2019.
- [8] G. Felix, G. Nápoles, R. Falcon, W. Froelich, K. Vanhoof, and R. Bello, "A review on methods and software for fuzzy cognitive maps," *Artificial Intelligence Review*, vol. 52, no. 3, pp. 1707–1737, 2019.
- [9] L. Ma and B. Sun, "Machine learning and AI in marketing - c," *International Journal of Research in Marketing*, vol. 37, no. 3, pp. 481–504, 2020.
- [10] K. Y. Chan, C. K. Kwong, P. Wongthongtham et al., "Affective design using machine learning: a survey and its prospect of conjoining big data," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 7, pp. 645–669, 2020.
- [11] S. Shafiqat, S. Kishwer, R. U. Rasool, J. Qadir, T. Amjad, and H. F. Ahmad, "Big data analytics enhanced healthcare systems: a review," *The Journal of Supercomputing*, vol. 76, no. 3, pp. 1754–1799, 2020.
- [12] M. I. Razzak, M. Imran, and G. Xu, "Big data analytics for preventive medicine," *Neural Computing & Applications*, vol. 32, no. 9, pp. 4417–4451, 2020.
- [13] M. M. B. Ismail, "A survey on content-based image retrieval [J]," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 5, pp. 159–170, 2017.
- [14] M. Avanzo, L. Wei, J. Stancanello et al., "Machine and deep learning methods for radiomics[J]," *Medical physics*, vol. 47, no. 5, pp. e185–e202, 2020.
- [15] N.-n. Zheng, Z.-y. Liu, P.-j. Ren et al., "Hybrid-augmented intelligence: collaboration and cognition," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 2, pp. 153–179, 2017.
- [16] A. Rahim Taleqani, J. Hough, and K. E. Nygard, "Public od bike sharing: a machine learning approach," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 4, pp. 195–204, 2019.
- [17] R. Kashyap and V. Tiwari, "Energy-based active contour method for image segmentation[J]," *International Journal of Electronic Healthcare*, vol. 9, no. 2-3, pp. 210–225, 2017.
- [18] L. Hou and R. J. Jiao, "Data-informed inverse design by product usage information: a review, framework and outlook," *Journal of Intelligent Manufacturing*, vol. 31, no. 3, pp. 529–552, 2020.
- [19] X. Chen, J. Qi, X. Zhu, X. Wang, and Z. Wang, "Unlabelled text mining methods based on two extension models of concept lattices," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 2, pp. 475–490, 2020.
- [20] Y. Li, Y. Zhuang, X. Hu et al., "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation[J]," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4035–4062, 2020.
- [21] A. Bonci, M. Pirani, and S. Longhi, "A database-centric framework for the modeling, simulation, and control of cyber-physical systems in the factory of the future," *Journal of Intelligent Systems*, vol. 27, no. 4, pp. 659–679, 2018.