IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 8, 2019, accepted October 22, 2019, date of publication October 28, 2019, date of current version November 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2949879

Analysis and Visualization Implementation of Medical Big Data Resource Sharing Mechanism Based on Deep Learning

YING YANG[1] AND TAO CHEN[2]

[1]School of Management, Xinxiang Medical University, Xinxiang 453003, China
[2]The Third Affiliated Hospital of Xinxiang Medical University, Xinxiang 453003, China

Corresponding author: Ying Yang (041121@xxmu.edu.cn)

ABSTRACT With the development of information technology, the informationization of the medical industry is also constantly developing rapidly, and medical data is growing exponentially. In the context of "Big Data +", people began to study the application of data visualization to medical data. Data visualization can make full use of the human sensory vision system to guide users through data analysis and present information hidden behind the data in an intuitive and easy-to-use manner. This paper first introduces the workflow of DBN, a deep learning algorithm, and summarizes the computational characteristics of the algorithm. The classification function is translated into an assembler using an instruction set-based assembly language, and the program is evaluated for performance. Secondly, based on the Hadoop ecosystem, this paper analyzes the BDMISS system for big data medical information resource sharing. Based on the system's requirements and functional positioning, from the medical information collection and sharing, data mining and knowledge management level, the big data medical service system is constructed. Based on the semantic network and ontology theory, big data mining technology and the design of "medical cloud", the resource sharing mechanism is analyzed. Based on the Spring MVC framework, using Echarts, HCharts and other data visualization technology, according to the design of specific modules, the visualization and display of medical data is realized, which has certain promotion effect on the research and development of medical big data visualization analysis.

INDEX TERMS Medical big data, data visualization, information resource sharing mechanism, deep learning.

I. INTRODUCTION

As the informationization of the medical industry continues to develop rapidly, medical data has grown exponentially, and medical big data has brought tremendous pressure on existing hospital information systems. With the emergence of various unstructured data, traditional medical information systems cannot meet the requirements of big data in terms of storage space, storage speed, storage structure, etc., some data is lost, resulting in loss of valuable medical data [1]–[3]. The data integrity of the system is not enough, and the data processing speed is slow, which cannot meet the user's demand for data visualization display [4]. In recent years, the information and communication industry of medical and health has ushered in the opportunity of development. The hospital information management system, public health service platform, telemedicine, mobile medical, and information equipment have formed a scale of 100 billion yuan. Medical informatization is no longer limited to transactional tasks such as hospital information management systems [5], [6]. Applications and research based onInternet of Things and cloud computing continue to deepen, and medical informatization begins to develop in the areas of process optimization and service innovation [7]. The Internet of Things and cloud computing have changed the patterns and paths of medical information services, optimized medical service processes, improved the efficiency of medical services, and transformed medical information sharing and service models [8], [9]. Deep learning not only improves people's lives in traditional industries, but as people's attention to physical health increases, machine learning algorithms closely related to deep learning have emerged in the health care field.

The associate editor coordinating the review of this manuscript and approving it for publication was Yongtao Hao.

Deep learning can learn about patient health data and medical records. These patient characterizations can be easily assisted in clinical testing, while deep learning can be embedded as a systematic framework in clinical decision making systems [10]. Relevant scholars have proposed a convolutional neural network to construct a ''medical record'' that can be used to help improve the theory of clinical diagnosis [11]–[13]. Prior to this, relevant scholars also described the construction of health characterization information through a sparse deep learning system to predict possible diseases [14], [15]. At present, the focus of visualization has shifted from ''scientific visualization'' focusing on scientific computing to the field of information data, and has formed a new research field ''information visualization'', which aims to heterogeneous, abstract and complicated data [16]. The collection of mining analysis work provides technical support [17]. Data visualization technology has evolved along with the continuous development of computer hardware technology, image processing technology and data mining technology, and its application scope has become more and more wide [18], [19]. From simple scientific computing field to medical science, geographic information science and product design, architectural design, fluid mechanics and other fields are closely related to information graphics, information visualization, scientific visualization and statistical graphics [20], [21]. At present, data visualization plays a crucial role in teaching research, program development and scientific research [22]. Researchers in these fields are also increasingly active in their research [23]. Data visualization has presented diversity at the technical level, and can be divided into some categories according to the principles of its use of technology [24]. Today's society is facing increasing pressure on medical needs and shrinking medical budgets [25], [26]. Data visualization as an important tool for data analysis not only plays an important role in hospital management, but also its powerful advantages are also reflected in limiting medical waste and control [27], [28]. The hospital has a wealth of data resources, including medical expenses data, electronic medical record data, medical imaging data, pathological parameter data, laboratory test data, etc. The visual requirements of these data make data visualization widely used in clinical and medical information systems [29]. As the scale of the graph continues to expand, there are characteristics such as rapid data changes, high complexity, and large data volume. The ability to store and traverse large-scale graphs has become increasingly important. The graph search algorithm has irregular memory access behavior, poor data locality, and low computational memory ratio. Although the parallel graph search algorithm has many years of research history, due to the high memory access calculation and the unbalanced computing load, the graph search algorithm is very inefficient and scalable in the existing high performance computing environment.

This paper studies the relevant theories and techniques of medical data visualization analysis. On the basis of this, combined with the needs of the health care big data platform, the needs of visual analysis of medical data (mainly resident health records and electronic information case data) are discussed, presenting valuable visual analysis requirements, and designing and visualizing the data using the appropriate appropriate charts, and constructing a visual analysis system based on medical big data. The workflow of DBN, a deep learning algorithm, is introduced, and the computational characteristics of the algorithm are summarized. The classification function is then translated into an assembler using an instruction set-based assembly language and the program is evaluated for performance. Based on the Hadoop ecosystem, a big data information sharing service system for regional medical informatization-BDMISS system is proposed. It discusses the system requirements, clarifies the positioning and functional subjects of the system, and clarifies the regional collaborative network and basic operational mechanism of the system. From the medical information integration and sharing, medical data mining and knowledge management, the BDMISS system was constructed, and detailed construction ideas at various levels were given. Based on the Spring MVC framework, the research uses Echarts, HCharts and other data visualization technology to realize the visual analysis and display of health medical data according to the design of specific modules, mainly the number of patients admitted to the hospital in 24H, age height and weight information, regional disease information, referral department information and medical examination information and other aspects.

The rest of this article is organized as follows. Section 2 discusses related theories and methods, followed by parallel optimization of deep learning algorithms for medical big data processing in Section 3. Section 4 analyzes the regional information resource sharing mechanism based on medical big data, and Section 5 realizes the visualization of medical big data. Finally, Section 6 summarizes the full paper and points out future research direction.

II. RELATED THEORIES AND METHODS

A. DEEP LEARNING NETWORK MODEL

This paper studies the DBN network structure of the deep learning algorithm, which consists of RBM (Restricted Boltzmann Machines) basic blocks. The structure of the deep learning tool chain is shown in Figure 1.

1) DBN NETWORK STRUCTURE

The DBN is a probability generation model consisting of multiple RBM layers. Figure 2 shows a DBN network structure consisting of three RBM networks. The training process is divided into Pretraining and Fine-tuning.

Pretraining: In order to retain more information, each layer of RBM network needs to be trained separately. The input to each layer of RBM is provided by the output of the previous layer of RBM. There is an encoding process and a decoding process for each block of the depth structure. Through the study of a large number of samples, the error between the original data x and the reconstructed data x' is continuously

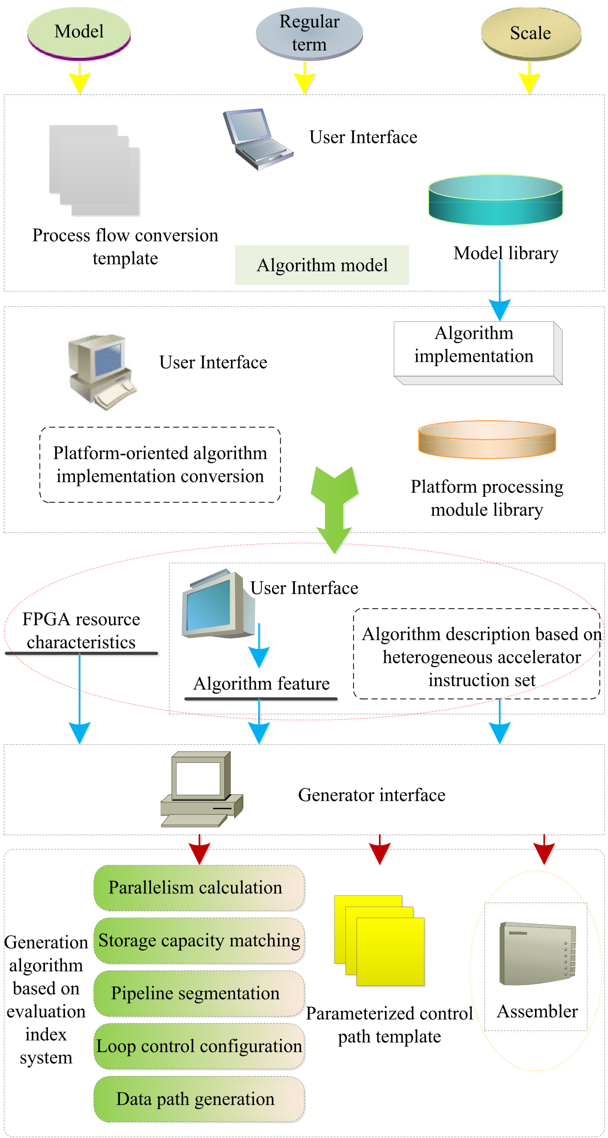


FIGURE 1. **Deep learning algorithm tool chain.**

reduced, thereby optimizing the network structure. For each layer, after RBM is in Pretraining, there will be a weight matrix W generated.

Fine-tuning: At the top of the RBM network, there is a tag data Label, which is used to evaluate the performance of the trained DBN network after the Pretraining is completed. For error backhaul, a backpropagation algorithm is used. In order to make the error meet the accuracy, it is necessary to fine tune a DBN network and modify the weight of the RBM network.

2) DBN TRAINING PROCESS

During the Pretraining process, multiple RBM networks perform unsupervised training one by one in stacking order. After the process is completed, the BP method of error control is used to perform fine adjustment of the weight.

The offset value is generally initialized to 0, and the weight matrix W is generally initialized with a Gaussian random function. The state of the hidden layer node is estimated

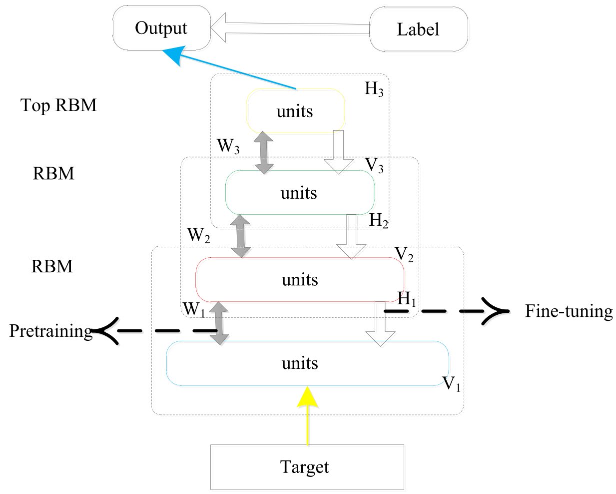


FIGURE 2. **DBN network structure.**

($v \to h$) using visible layer nodes. We calculate the probability values for each hidden layer node with a value of 1.

The weights and offset values used are w_{ji} and b_j, respectively, and the multiplication of the value v_i of each node and the corresponding connection weight w_{ji} are added, and the sum is added to the offset value b_j of h_j, and then a sigmoid function is applied, that is, $f(x) = 1/\exp(-x)$. The value of $p(h_j = 1|v)$ represents the activity of h_j under the action of each visible layer node, and its calculation formula is as follows:

$$p(h_1 = 1\,|V) = sigmoid(\sum_{i=1}^{D} v_i w_{1i} + b_1) \tag{1}$$

Compared with the formula (1), in the formula (2), the weight matrix is subjected to one transposition, and the offset value used is the offset value a_i of the i-th visible layer node.

$$p(h_j = 1\,|V) = sigmoid(\sum_{i=1}^{D} v_i w_j + b_j) \tag{2}$$

The update rules for the network parameters (weight matrix W and offset values a and b are as follows, where ε is the learning rate.

$$w'_{i,j} = w_{i,j} + \varepsilon(v_i h_j - v'_i h'_j) \tag{3}$$
$$a'_i = a_i + \varepsilon(v_i - v'_i) \tag{4}$$
$$b'_j = b_j + \varepsilon(h_j - h'_j) \tag{5}$$

B. MEDICAL BIG DATA TECHNOLOGY

The valuable medical data resources have invaluable value for predictive management and control of disease, medical research, and medical informationization. With the promotion of cloud computing, medical informationization is bound to develop unprecedentedly, and the era of medical big data is bound to come. Of course, cloud computing is the most effective means of analyzing and utilizing big data while generating big data.

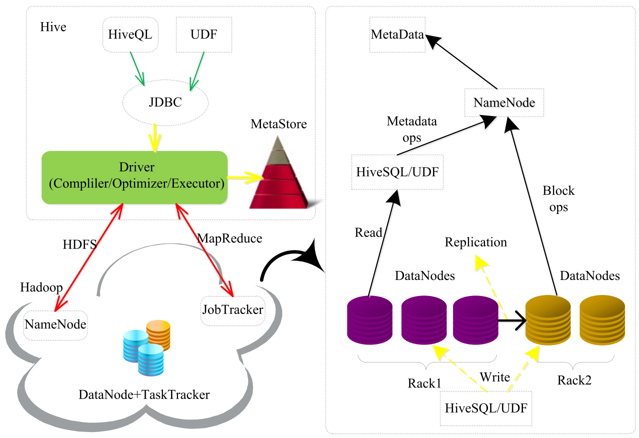


FIGURE 3. Medical big data warehouse.

In terms of clinical operations, there are five major scenarios for big data applications. According to McKinsey estimates, the full use of such big data applications will significantly reduce medical expenses.

(1) In the aspect of comparative effect research, the patient's personal characteristic information data is comprehensively compared and analyzed, and then various treatment measures are deeply compared to finally determine the best treatment plan.

(2) The clinical decision support system uses a data-driven clinical decision support system to make itself smarter by using big data analysis technology, which can improve the efficiency of medical workers and the quality of medical services.

(3) Visualization of medical big data enhances the transparency of medical data and processes. It can promote the optimization of medical business processes, reduce the cost of medical care and improve the quality of medical services.

(4) Through remote monitoring and recording of data on chronically ill patients through a variety of wearable health devices, the collection and analysis of big data can help medical workers develop treatments for the patient.

(5) Advanced analysis of patient cases and big data analysis of patient files can predict the patient's susceptibility to various diseases.

The medical big data warehouse is shown in Figure 3. The establishment of the big data storage and management module meets the requirements of the data warehouse, which is a subject-oriented, integrated data warehouse based on time variables and decision support.

C. DEEP LEARNING DISTRIBUTED TRAINING METHOD UNDER MEDICAL BIG DATA

Traditional distributed training is generally a CPU, which is very convenient in dealing with general big data problems such as text calculation and log processing. However, the big data processing training under the deep learning of this study is very different. Firstly, deep learning involves a large number of iterative operations. These operations are interdependent and difficult to be independent. Therefore, in order to accelerate, it is necessary to use top-level GPU parallel acceleration. It speeds up matrix operations and allows multiple machines to be inserted into a single card.

Data parallelism is to evenly distribute data to multiple training nodes for training at the same time of training. The parameters are synchronized according to a certain period. The model copy on each training node is responsible for a part of the data training task.

Data parallelism can be highly fault-tolerant, and the entire training will not stop because a node hangs. This is especially important in tens of thousands of nodes in large companies like Google. An improvement to this asynchronous parallel gradient algorithm proposed in the early years of data parallelism is called Downpour SGD.

First, you need to transfer the data slices to the corresponding training nodes. This process is similar to the traditional big data platform. These data can also be distributed storage and stored in a distributed file system. Then each data node uses a copy of the model copy separately, so that the training is independent of each other, and the hardware resources are allocated to achieve horizontal expansion. This kind of structure can make several modifications to the existing framework, such as minimum modification, efficiency compliance, and convenient expansion. The role of the parameter server here is to coordinate all data exchange and data update, and it is also the performance bottleneck of the entire framework. Therefore, how to design the data update algorithm of this parameter server is very important to the performance of the whole system.

Parameter server communication has the distinction between synchronous and asynchronous. Here Downpour SGD is an asynchronous gradient update algorithm: it has two important parameters nfetch and npush. The role of this parameter is as the name implies, fetch means to take parameters from the parameter server. This process is when a training node has just completed a small iteration operation and needs to make a request to the parameter server when the next calculation is needed. The frequency of initiating a request is the fetch control of this parameter. Then it pushes parameter, which means that the training node will actively push the current gradient difference to the parameter server, so that the parameter of the parameter server can be updated in time to achieve the coordination function. This coordination frequency is controlled by this parameter. The size of these two parameters is the same, and has a good update effect on the network, but it can't be too big or too small. Too large a result causes the expiration parameters to be too long, and the updated gradient will cause the network convergence to fluctuate very much. Too small, this architecture will gradually approach the synchronous update, and the network communication requirements are very high.

D. DATA VISUALIZATION TECHNOLOGY

The main point of big data display technology lies in: the result information from the statistical analysis of big data is displayed to the user in an intuitive way, and it is convenient for the user to understand while displaying complex data information. For large Internet companies, since most of the work is carried out around the website, the technology for front-end development is relatively mature, so it will establish its own data display R&D team. However, for some small and micro enterprises, the cost of setting up and maintaining a data display R&D team is relatively high. However, due to the rapid development of open source software in the modern era, large companies such as Baidu have supported open source data display technology. Because big data visualization technology mainly relies on web page display, Java Script and HTML5 tags are still the necessary foundation. Currently, the open source front-end display frameworks are D3.js, Google Chart, High Charts and Echarts.

High Charts is a chart library. Because it is written in pure Java Script, it can quickly and easily add various charts with user interactions on the web site or application to better satisfy the user experience.

ECharts open source from Baidu business front-end data visualization team, based on html5 Canvas, is a pure Javascript chart library, providing data visualization charts that are intuitive, vivid, interactive, and customizable. Innovative drag-and-drop computing, data views, and range roaming features greatly enhance the user experience and give users the ability to mine and integrate data.

From the perspective of big data analysis, it plays an important role in the visualization of big data in information analysis technology. Different types of information from data can be divided into two categories, namely spatiotemporal data and non-spatiotemporal data. Spatio-temporal data represents data with geographic location and time stamps.

III. PARALLEL OPTIMIZATION OF DEEP LEARNING ALGORITHMS FOR MEDICAL BIG DATA PROCESSING

A. FINE_TUNING PROCESS ANALYSIS

The Fine-tuning process is implemented as follows:

Step1: Initialize parameters, randomly initialize the parameters of the top-level RBM back propagation, set the learning rate, training step size, maximum number of iterations, and so on.

Step2: Calculation of the forward transfer process. For the input samples, the output values of the RBMs of each layer are calculated, and finally the output values of the DBN network are obtained. The value of the tag data is subtracted from the DBN network output value to obtain an error value.

Step3: Error backpropagation, from the output layer to the input layer, calculate the equivalent error of each layer, and correct the network weight parameter.

Step4: Adjust the weight matrix between layers according to the learning rate.

Step 5: Determine whether the accuracy requirement or the number of iterations is met. If yes, return to Step 2 and perform a forward calculation based on the recalibrated weight; otherwise, the training ends.

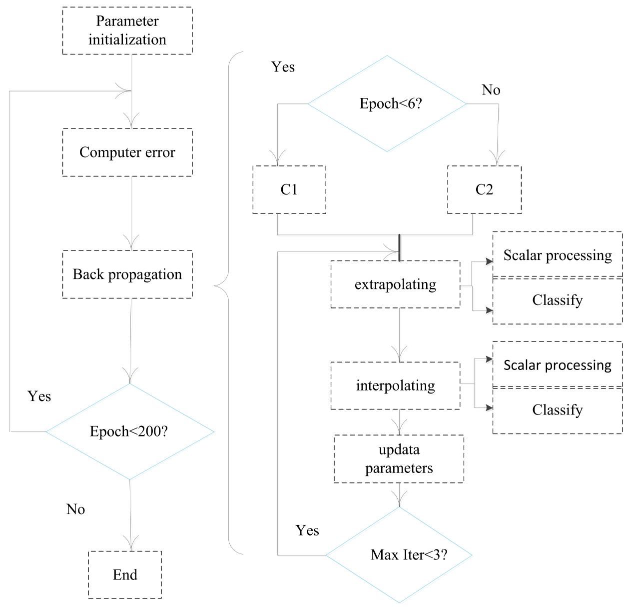


FIGURE 4. Fine_tuning parallel optimization process.

This paper performs parallel optimization for the Fine_tuning process. The specific execution process of Fine_tuning is as follows:

Figure 4 shows the execution of Fine_tuning. By analyzing the DBN source code, the Fine_tuning process is divided into three parts: Parameter Initialization, Compute Classify Error, and Back Propagation.

The backpropagation process performs a reverse backhaul based on the error value, adjusting the value of the RBM weight matrix to minimize the error value. As shown in Figure 4, the minimum error is obtained by performing Linesearch three times. Each linear search requires Conjugate Gradient to determine the search direction and step size. When the number of iterations is less than 6, the weight matrix of the topmost RBM is fine-tuned according to the result of function 1 (C1), and the RBM weights of other layers remain unchanged; when the number of iterations is greater than or equal to 6, according to the classifier function 2, the classification result of C2 is fine-tuned for the RBM of all layers.

The linear search uses the conjugate gradient to calculate the direction and step size of the search, by using the low-order (usually no more than three) polynomial to approximate the objective function in the search interval, and using the minimum point of the interpolation polynomial to approximate the minimum value of the objective function.

B. FUNCTION PARALLEL OPTIMIZATION CODE VERIFICATION

This paper mainly encodes and verifies the two classification functions in the Fine_tuning process. Considering that the FPGA hardware accelerates the development cycle, it is not easy to debug. This article first encodes the classification function code using assembly instructions, and then uses the

TABLE 1. Super vector instructions.

Name	Code	Des.	Source1	Source2	Size
VLOAD	80	Ri	Rj	/	Size
VSTORE	81	Ri	Rj	/	Size
VVMC	41	Ri	Rj	Rk	Size
VMAC	40	Ri	Rj	Rk	Size
VACC	23	Ri	Rj	Imm	Size
VADD	20	Ri	Rj	Rk	Size
VSUB	21	Ri	Rj	Rk	Size
VDIV	24	Ri	Rj	Imm	Size
VSMUL	25	Ri	Rj	Imm	Size
VSADD	22	Ri	Rj	/	Size
VSIG	30	Ri	Rj	/	Size
VSTAS	31	Ri	Rj	/	Size

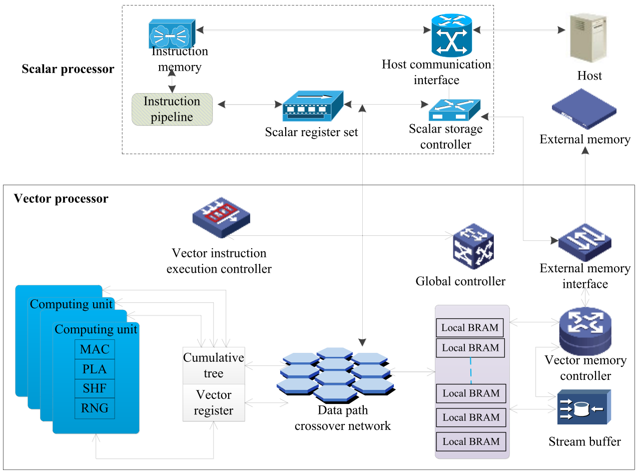


FIGURE 5. Deep learning accelerator framework.

assembler simulator to verify the correctness. The assembly code design takes into account the FPGA on-chip storage resources and DSP computing resources, so as to reasonably set the number of matrix basic operation blocks and arithmetic units. Considering that the matrix operation in the classification function is an operation of a dense matrix, it is different from the sparse matrix operation in the graph search. Therefore, we can take the data in the external DRAM to the on-chip cache for calculation, and make full use of the locality principle to improve the performance of the algorithm accelerator. For matrix operations, the supervector instruction encoding in Table 1 is used.

Table 1 shows the super vector instructions set for machine learning features. Each instruction is divided into 5 functional segments, namely code segment, destination address (Des.), source address 1 (Source1), and source address 2 (Source2), data block size (Size). For unneeded function segments, the default setting is 0.

Figure 5 shows the accelerator framework for deep learning. The accelerator is divided into two parts, the scalar processor and the vector processor.

The scalar processor consists of a Host communication interface, instruction memory, instruction pipeline, variable register banks, and a scalar memory controller. The host completes the configuration of the scalar processor through the Host communication interface. The scalar processor mainly handles some global shared variables or constants.

The vector processor consists of a vector instruction execution controller, a global load/store controller, a data path crossover network, a vector register, an accumulation tree, a vector storage controller, and an external memory interface. It mainly completes the operation of the vector to achieve parallelization of the calculation. The vector processor uses a global Load/Store controller to read the data stored externally from the external memory interface into the on-chip buffer of the FPGA. The data is written to the on-chip local BRAM by the stream buffer control unit under the control of the vector memory. The data required by multiple computing units at the time of the computation is obtained from the local BRAM or the scalar register bank via the data path crossover network. The various basic arithmetic units required for the calculation unit, such as matrix multiplication, matrix addition, etc., are selected by the bypass switch.

When the program is completely pipelined, each scalar instruction takes one clock cycle. The execution time of each super vector instruction is related to the basic matrix block size. The execution time of each fetch instruction is related to the storage control.

Here, we use the DDR2 memory controller, each time we need to fetch the matrix block to the on-chip memory or save it back to the external DRAM. It is found that the memory access time is related to the size of the operation block of the matrix. For example, if the matrix block time of size M is T, the time of the matrix block of size N is NT/M.

The use of parallel-optimized classification functions yielded significant performance improvements over the CPU platform. Since the processing function is the basic processing function of the Fine_tuning process, it needs to be called frequently, and the time required for the Fine_tuning process takes up most of the DBN network training time (more than 90%). Therefore, parallel optimization of the classification function can reduce the training time of the DBN network, thereby reducing the training time of deep learning and improving the overall execution efficiency of the program.

IV. ANALYSIS OF REGIONAL INFORMATION RESOURCE SHARING MECHANISM BASED ON MEDICAL BIG DATA

A. IoT-BASED MEDICAL DATA COLLECTION

The medical detection mainly implements the electronic instrument for data detection and uploads the detected data to the access gateway through the sensor network, thereby transmitting the detection data to the server through the network, and can store and intelligently analyze the detected data to reach the data. The purpose of detection, data accumulation, and intelligent data analysis is to achieve the data base accumulation and service functions of smart medical care.

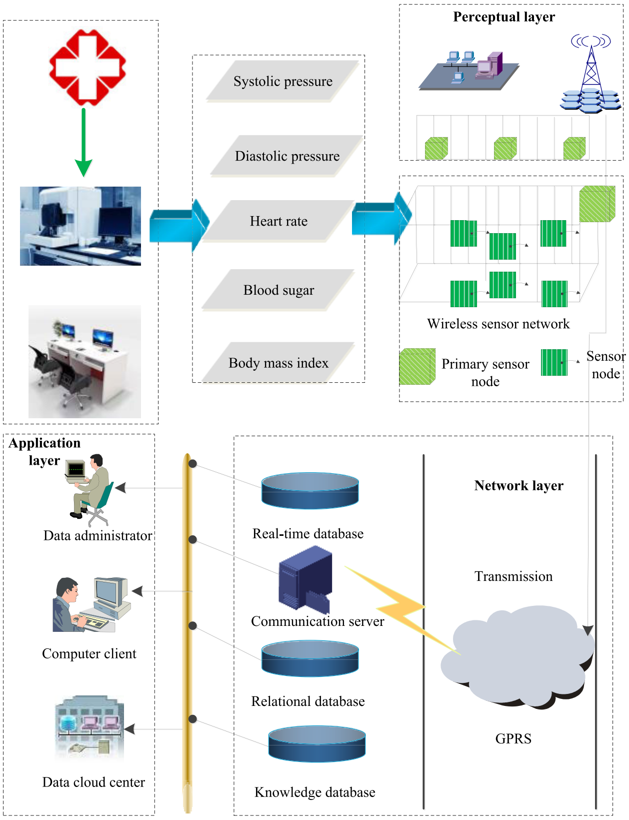


FIGURE 6. Schematic diagram of IoT data acquisition network architecture.

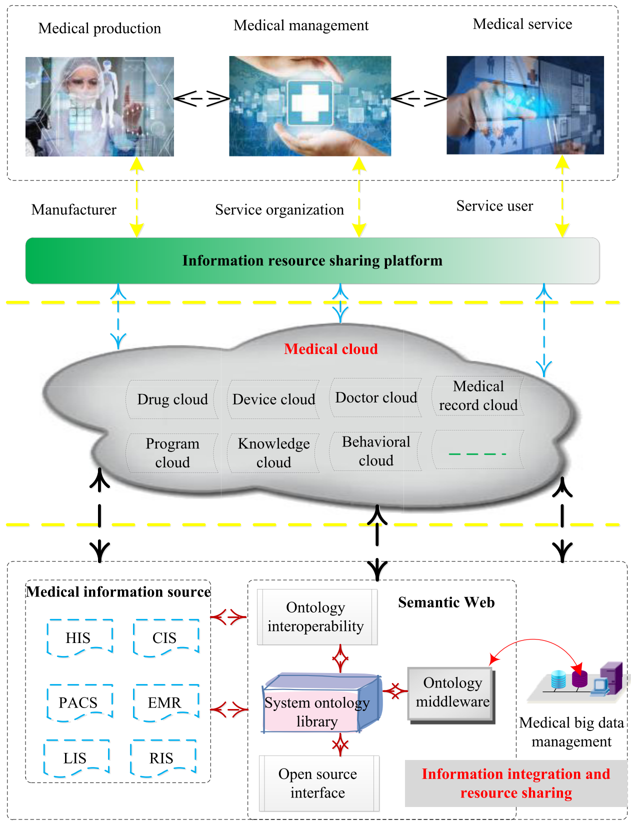


FIGURE 7. BDMISS architecture.

The health test is mainly described by taking the blood pressure test as an example.

Firstly, the blood pressure meter should be connected with the intelligent terminal (PAD, smart phone, etc.) through the Bluetooth protocol, and realized by mac addressing; after the connection is established, the Socket pipeline is established to realize the pipeline connection; after the pipeline connection is established, the blood pressure meter sends the detection result. For the intelligent terminal, the result is included in the data protocol of the blood pressure detecting device; the intelligent terminal receives the data protocol, analyzes the data protocol, obtains the blood pressure detection result, and performs data verification to realize the detection data to the access gateway. Finally, the access gateway uploads the data to the server through the network layer and stores it in the database. The electronic instrument test transmits the detected result to the intelligent terminal for receiving and parsing. The Internet of Things data acquisition network architecture is shown in Figure 6.

After the personal data in the health check is uploaded to the server via Bluetooth and the network, it will be stored in the personal data table in the database, and all the personal health data will be stored, allowing the user to query all the personal health data and give the query. In addition, the user is allowed to delete abnormal data (abnormal data due to improper user operation).

In addition, trend analysis of the user's health data based on the data information known to the user enables the user to understand their own data trends and adjust the corresponding habits and living conditions.

B. MEDICAL INFORMATION RESOURCE INTEGRATION AND SHARING

Information integration and sharing is the foundation and support link of the big data medical information sharing system. Regional medical collaborative network structure and information sharing mechanism are important guarantees for the accuracy and availability of medical information, and have a profound impact and decisive role in the efficiency and quality of subsequent data mining and medical knowledge management. The BDMISS system requires an efficient and highly stable infrastructure and organizational structure to provide a stable foundation for the entire system. The BDMISS architecture is shown in Figure 7.

Based on the regional medical information service platform, the medical information is managed and integrated in a unified manner, and the regional medical "medical cloud" is constructed to realize the efficient integration and knowledge discovery of medical information. Through the BDMISS service platform, information sharing and accurate medical information services for medical demanders can be realized, and the precise services of O2P2O mode for online and offline resources and system users can be realized. The overall architecture of the system is divided into three levels: information integration and sharing, data mining and knowledge management, and medical services.

Medical management is a government-led management model. From the top-level design of the health care sector to the health care management system, government departments have a leading decisive role. To analyze the collaborative network and operational framework for regional medical informatization, it is necessary to form a unified service mechanism and guarantee system with the government department as the core, and realize the clear division of labor and functional coordination of regional medical institutions. The collaborative network of the BDMISS system is based on the government and the departments directly under the medical and health management department. It develops a unified information sharing mechanism and guarantee system, and integrates medical service organizations in the region into a regional medical service community. A complete management chain and operational structure are formed by connecting demanders and suppliers of medical services.

The core of the operation and management of the system is to realize the integration and sharing of regional medical data and resources. The system adopts Hadoop HDFS system to realize information and data integration and unified management of various medical service organizations in the region, and realize the sharing of medical data, medical equipment, medical knowledge and other resources, thereby realizing universal medical care, promoting medical fairness and improving medical security. HDFS (Hadoop Distributed File System) is an open source project of Apache. HDFS can realize PB level data set processing, provide high throughput, high reliability data sequence access, and realize computing node sharing with low cost hardware facilities.

This paper uses the characteristics of HDFS to realize the integration and sharing of cross-platform regional medical information, centrally manage the data of each medical service information system, realize information sharing and resource integration, and build a system knowledge base and resource pool. According to the specific needs and preferences of users, through the big data analysis system to achieve accurate recommendation of medical services and intelligent matching of supply and demand, to achieve on-demand distribution of regional medical resources, BDMISS information sharing and service docking mode is shown in Figure 8.

The problem of benefit distribution is an important guarantee mechanism for system operation and collaborative service. The realization of BDMISS system requires a perfect benefit distribution mechanism. The issue of benefit distribution in the medical service system is a problem in the field of financing in the medical delivery system. From the current international operation, the main sources of financing include medical expenses paid by patients, social insurance fees, government taxes and other income. The income source of the medical service system can be mainly divided into two aspects: one is the service fee for providing medical services, including the part paid by the patient and the part of the medical insurance; on the other hand, it is the financial subsidies and policy subsidies provided by the government.

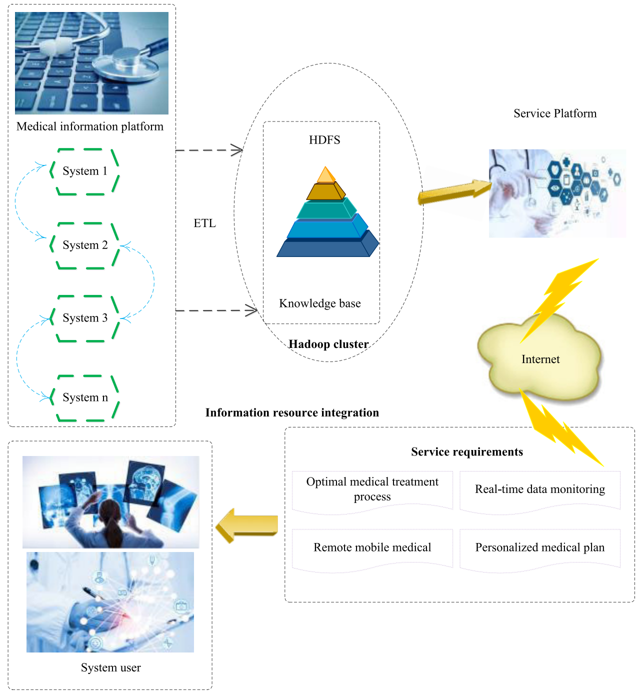


FIGURE 8. Information sharing and service docking mode.

The BDMISS system constructed in this paper uses the medical information platform of the regional medical service organization as the information source, and adopts the hybrid ontology method to construct the semantic model, that is, constructs the global ontology for the whole service system, and each data source defines the corresponding ontology to realize the semantics. Each data source constructs its own local ontology, and at the same time constructs a global ontology for the whole service system, sharing the technical terminology and conceptual model of the system, introducing a mature biomedical ontology library as the upper ontology, and solving the terminology and concept mismatch between the system ontology. The local ontology is mapped to the global ontology, and the information resources are shared by using the ontology middleware such as Jena.

C. MEDICAL BIG DATA INFORMATION RESOURCE MINING AND KNOWLEDGE SHARING

The BDMISS system information sharing is to promote the social development of medical knowledge and services. This paper combines external medical information and knowledge base to build a data mining and knowledge management system supported by Hadoop data analysis and processing technology, and finally realizes the cloud service model of resource sharing with "medical cloud" architecture as the core.

The system architecture of Hadoop Map Reduce is shown in Figure 9.

Based on the Semantic Web's medical information integration and sharing architecture, this paper builds medical data mining and knowledge management based on data center network, big data storage, big data processing, cloud computing

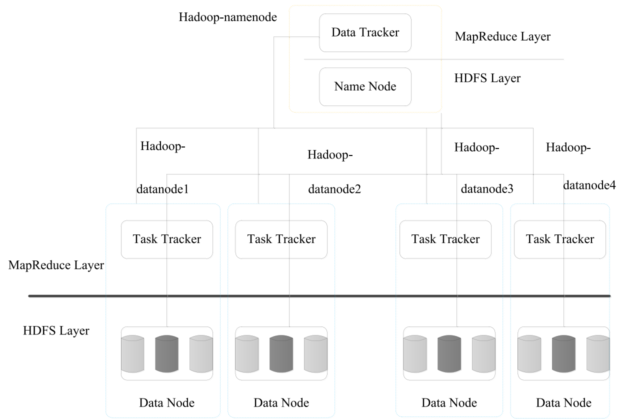


FIGURE 9. System architecture of hadoop map reduce.

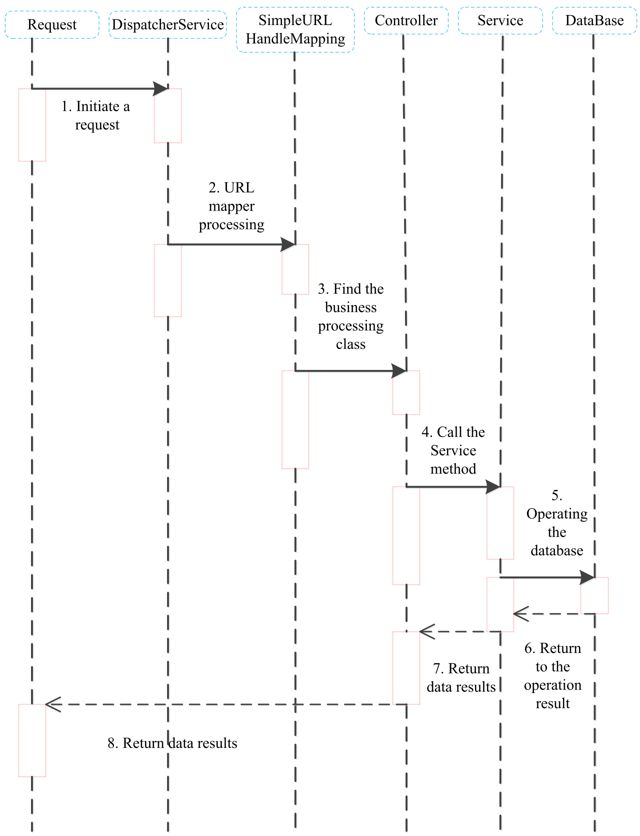


FIGURE 10. Spring MVC request timing diagram.

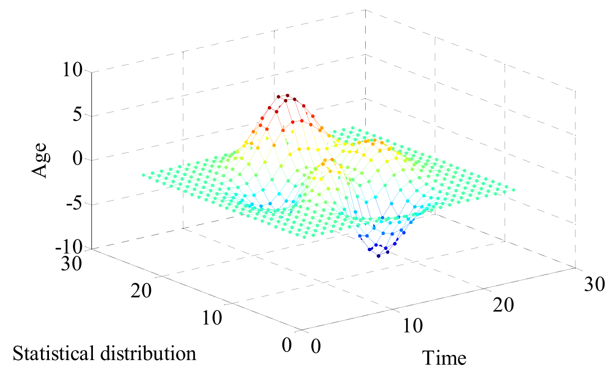


FIGURE 11. Statistical analysis of the age of patients discharged from hospital in 24H.

and machine learning by using the parallel processing capability of Hadoop ecosystem and Map Reduce.

V. MEDICAL BIG DATA VISUALIZATION IMPLEMENTATION

A. IMPLEMENTATION OF THE BASIC FRAMEWORK

The medical big data-based visual analysis system is developed based on the Java EE platform, using the open source lightweight framework Spring, and using Spring MVC in Spring Frame Work to build the MVC architecture, using Hibernate as the data persistence layer framework. The basic timing diagram of the requested process is shown in Figure 10.

In the system development process, the Model and View are combined at the same level, interacting with the Controller layer through the HTTP protocol, and the information is passed as a JSON string. The Controller calls the Service to process the business logic, and obtains the processing result and message from the Service as org.json.JSONObject. When the business logic needs to operate the database, the Service calls the interface provided by the DAO layer, and the DAO layer interacts with the database through Hibernate.

Maven is used for project management during development. By modifying the Maven configuration, it is easy to download the required packages from Maven's central repository. The IDE used for development is the Eclipse Mars.2 Release, which has been integrated with the Maven plugin to make it easy to create Maven projects.

We create a new Maven project, add some basic dependencies of the java web project and framework dependencies to be used in the Maven configuration file, such as Spring, Hibernate, etc., and then add some useful tools, such as json, log4j, junit, etc. Weka, etc., to complete the construction of the development framework.

B. IMPLEMENTATION OF MEDICAL DATA VISUALIZATION PART MODULE

1) AGE INFORMATION OF PATIENTS ADMITTED TO 24H

The user clicks on the "24H Into and Discharge Patient Information" statistical analysis in the drop-down menu of the menu bar, the system sends the request to Flask. After receiving the request, Flask sends the relevant data application for calculating the patient within 24 hours to the SPARK platform. SPARK platform calculates the parameter values needed to draw the age chart of the patient who is discharged from the hospital, and stores the calculated data in the My SQL database. Flask obtains the data (age and number) required for the chart from the My SQL database. After formatting, it is sent to the browser in JSON format. The front end analyzes the data and displays the statistical graph of the number of patients discharged from the hospital in each age group. The visualization page is shown in Figure 11.

The user clicks on the "Hospital Expense Information" statistical analysis in the drop-down menu of the menu bar, and the system sends the request to Flask. After receiving the

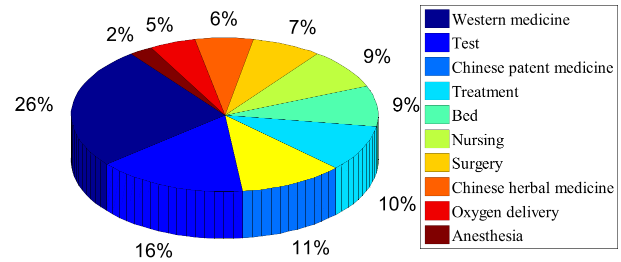


FIGURE 12. Pie chart of hospitalization expenses statistical analysis.

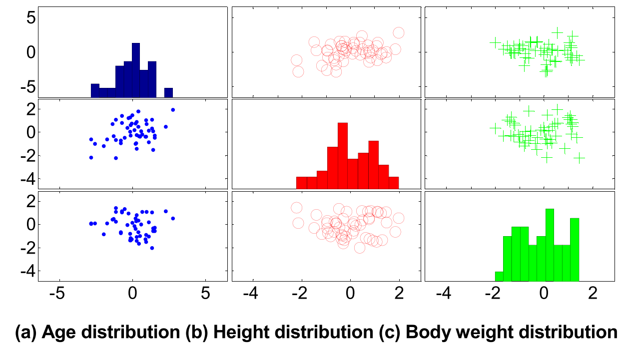


FIGURE 13. Statistic analysis of age height and weight statistics.

request, Flask sends the relevant data request for calculating the hospitalization expense information to the SPARK platform, and the SPARK platform calculates the hospitalization expense information. The parameter values required for the chart are stored, and the calculated data is stored in the My SQL database. Flask obtains the data (category and cost data) required for the chart from the My SQL database, and formats it in JSON format. To the browser, the front end analyzes the data and displays the pie chart of the hospitalization expense information on the display interface. The data visualization page for the hospitalization expense information is shown in Figure 12.

2) AGE HEIGHT AND WEIGHT INFORMATION

The user selects the statistical analysis of the ''age height and weight information'' of the examiner in the drop-down menu of the menu bar, and the system sends the request to Flask. After receiving the request, Flask sends the relevant data request for calculating the hospitalization fee information to the SPARK platform, and the SPARK platform calculates and draws. The calculated data is stored in the My SQL database. Flask obtains the data (age, height, weight) required for the medical examiner information chart from the My SQL database. After being sent to the browser in JSON format, the front end analyzes the data and displays the age-height weight scatter plot of the examiner on the display interface. Each point represents a person's related information, and the point-intensive places indicate a large number of people. The data visualization page for age height and weight information is shown in Figure 13.

3) REGIONAL DISEASE INFORMATION

The user clicks on the ''Sickness Information'' analysis in the drop-down menu of the ''Public Medical Information

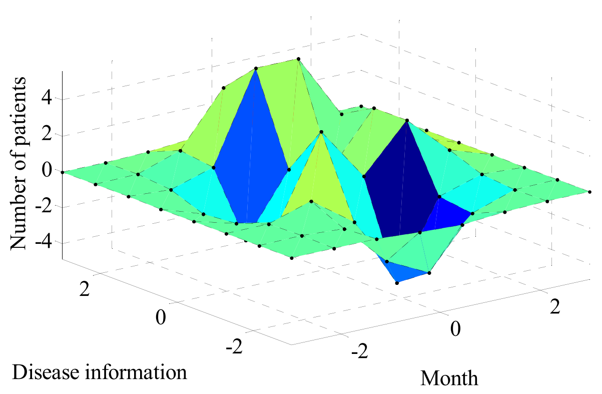


FIGURE 14. Visualization of disease information data.

Service'' in the menu bar. The interface will first appear in the interface, the user can click on the ''town'' that he wants to view, and then he will jump to the interface of the statistical information display. The display box will display the ''Year'' and ''Disease Type'' selection boxes. The user can click on the year and disease type information that he wants to view. The system sends the request to Flask. After receiving the request, Flask sends the application for calculating the disease drug association job to the Spark platform. On the Spark platform, the parameter values required to draw the ''by-town disease information'' chart are calculated, and the calculated data is stored in the My Sql database. Flask obtains the ''by town disease information'' chart from the My Sql database. The required data is formatted and sent to the browser in JSON format. The front end displays the trend graph of the number of people in the corresponding year and disease type on the display interface by analyzing the data. The data visualization page for regional disease information is shown in Figure 14.

4) MEDICAL EXAMINATION INFORMATION

The user clicks on ''Medical Analysis'' in the drop-down menu of ''Regional Medical Service'' in the menu bar, and the system sends the request to Flask. After receiving the request, Flask sends the calculation medical examination analysis job application to the SPARK computing platform, and the SPARK platform calculates the physical examination analysis chart. The required parameter values and the calculated data are stored in the My SQL database. Flask obtains the physical analysis data from the My SQL database, formats it and sends it to the browser in JSON format. The front end parses the data and displays it. The physical examination analysis chart is displayed on the interface, and the data visualization page for the medical examination information is shown in Figure 15.

5) REFERRAL DEPARTMENT INFORMATION

The user clicks on the ''Review Department Statistics'' analysis in the drop-down menu of the ''Public Medical Information Service'' in the menu bar, and the system sends the request to Flask. After receiving the request, Flask sends the application for calculating the disease drug association

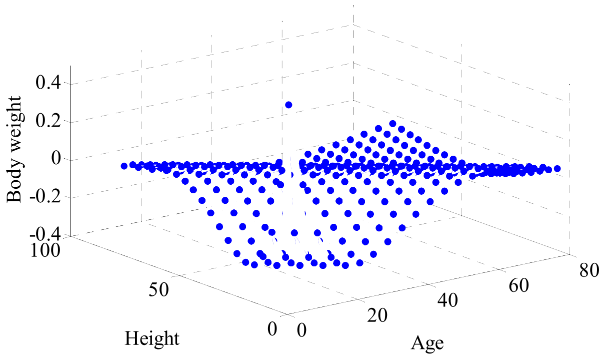


FIGURE 15. **Physical examination analysis.**

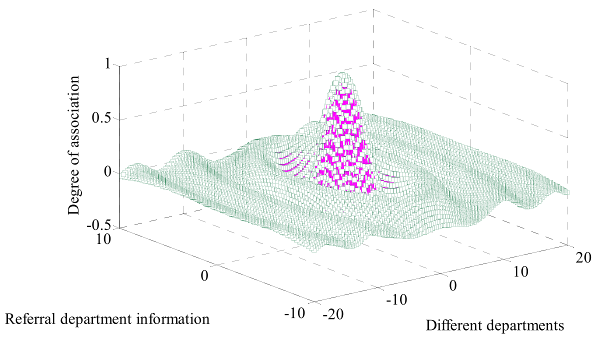


FIGURE 16. **Referral department association diagram.**

job to the SPARK platform, and the SPARK platform calculates Draw the parameter values required for the "referral department association map" and store the calculated data in the My SQL database. Flask obtains the referral department correlation data from the My SQL database, and formats it in JSON format. The front end analyzes the data and displays the referral department information association diagram on the display interface. The different colored squares in the figure represent different departments, and the connection indicates the situation of transferring out and transferring to the department. Putting the mouse over the line or square will display more detailed information. The referral map of the referral department is shown in Figure 16.

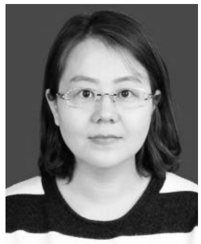
VI. CONCLUSION

This paper aims to study the use of visualization technology to better serve the medical big data platform, allowing users to participate subjectively throughout the visual analysis process. This paper analyzes the workflow of DBN, a deep learning algorithm, and summarizes the operational characteristics of the algorithm. The DBN's classification function was translated into an assembler using an instruction set-based assembly language, and the program was evaluated for performance. For the regional medical information resources, the BDMISS is proposed. From the actual needs and applications, the requirements, overall architecture, functional positioning, functional subjects and operational mechanisms of BDMISS are clarified. Based on the construction ideas of medical information integration and sharing, information mining and knowledge management, the implementation plan of BDMISS is formed. Compared with the traditional medical service system, the BDMISS system has obvious progress and superiority in terms of service ability, management ability and innovation ability. The research results of this paper have far-reaching significance and reference value for medical reform and development. In the medical data visualization analysis, the "24H outpatient admission age information", "outpatient cost information", "age height and weight information", "regional disease information", etc. are mainly realized. However, for medical data information, the scale will gradually increase, the types will gradually increase, and the structure will become more and more complex. So, there are still a lot of valuable information to be mined and analyzed. As the demand of users increases and the amount of medical data increases, the functional requirements in the medical data visualization system will increase accordingly. For the unfinished functions, further research and development is needed to better serve the residents' medical care and improve the information sharing of medical resources.

REFERENCES

[1] W.-J. Kim, "Knowledge-based diagnosis and prediction using big data and deep learning in precision medicine," *Investigative Clin. Urol.*, vol. 59, no. 2, pp. 69–71, 2018.
[2] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, and E. I.-C. Chang, "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC Bioinformatics*, vol. 18, no. 1, 2017, Art. no. 281.
[3] S. R. Panta, R. Wang, J. Fries, R. Kalyanam, N. Speer, M. Banich, K. Kiehl, M. King, M. Milham, T. D. Wager, J. A. Turner, S. M. Plis, and V. D. Calhoun, "A tool for interactive data visualization: Application to over 10,000 brain imaging and phantom MRI data sets," *Frontiers Neuroinformatics*, vol. 10, no. 11, p. 9, 2016.
[4] H. Liu, T. Taniguchi, Y. Tanaka, K. Takenaka, and T. Bando, "Visualization of driving behavior based on hidden feature extraction by using deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2477–2489, Sep. 2017.
[5] Y. N. Zhou, W. Zhao, and Y. N. Fan, "Real-time rendering and interactive visualization of remote sensing big data," *J. Geoinf. Sci.*, vol. 18, no. 5, pp. 664–672, 2016.
[6] H.-C. Shin, L. Lu, L. Kim, A. Seff, J. Yao, and R. M. Summers, "Interleaved text/image Deep Mining on a large-scale radiology database," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1–31, 2016.
[7] D. V. Dimitrov, "Medical Internet of Things and big data in healthcare," *Healthcare Inform. Res.*, vol. 22, no. 3, pp. 156–163, 2016.
[8] K. Rall, M. L. Satterthwaite, A. V. Pandey, J. Emerson, J. Boy, O. Nov, and E. Bertini, "Data visualization for human rights advocacy," *J. Hum. Rights Pract.*, vol. 8, no. 2, pp. 171–197, 2016.
[9] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
[10] J. Zhang and M. L. Huang, "Density approach: A new model for bigdata analysis and visualization," *Concurrency Comput. Pract. Exper.*, vol. 28, no. 3, pp. 661–673, 2016.
[11] X. Fan, Q. Yao, Y. Cai, F. Miao, F. Sun, and Y. Li, "Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 6, pp. 1744–1753, Nov. 2018.
[12] S. Hartong, "Between assessments, digital technologies and big data: The growing influence of 'hidden' data mediators in education," *Eur. Educ. Res. J.*, vol. 15, no. 5, pp. 523–536, 2016.
[13] Y. Yamamoto, "Healthcare and the roles of the medical profession in the big data era," *Jpn. Med. Assoc. J.*, vol. 59, nos. 2–3, pp. 125–139, 2016.
[14] M. Liu, J. Shi, K. Cao, J. Zhu, and S. Liu, "Analyzing the training processes of deep generative models," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 77–87, Jan. 2018.

[15] N. K. Kasabov, M. G. Doborjeh, and Z. G. Doborjeh, "Mapping, learning, visualization, classification, and understanding of fMRI data in the NeuCube evolving spatiotemporal data machine of spiking neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 4, pp. 887–899, Apr. 2017.
[16] J. P. Wanderer, S. E. Nelson, J. M. Ehrenfeld, S. Monahan, and S. Park, "Clinical data visualization: The current state and future needs," *J. Med. Syst.*, vol. 40, no. 12, p. 275, 2016.
[17] A. Dasgupta, J. Poco, E. Bertini, and C. T. Silva, "Reducing the analytical bottleneck for domain scientists: Lessons from a climate data visualization case study," *Comput. Sci. Eng.*, vol. 18, no. 1, pp. 92–100, 2016.
[18] J. Wang, L. Gou, H. Yang, and H.-W. Shen, "GANViz: A visual analytics approach to understand the adversarial game," *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 6, pp. 1905–1917, Jun. 2018.
[19] T. J. Brigham, "Feast for the eyes: An introduction to data visualization," *Med. Reference Services Quart.*, vol. 35, no. 2, pp. 215–223, 2016.
[20] J. Zhu, A.-A. Liu, M. Chen, T. Tasdizen, and H. Su, "Special issue on biomedical big data: Understanding, learning and applications," *IEEE Trans. Big Data*, vol. 3, no. 4, pp. 375–377, Dec. 2017.
[21] G. Xiong, F. Zhu, X. Dong, H. Fan, B. Hu, Q. Kong, W. Kang, and T. Teng, "A kind of novel ITS based on space-air-ground big-data," *IEEE Intell. Transp. Syst. Mag.*, vol. 8, no. 1, pp. 10–22, Spring 2016.
[22] S. Sohangir, D. Wang, A. Pomeranets, and T. M. Khoshgoftaar, "Big Data: Deep Learning for financial sentiment analysis," *J. Big Data*, vol. 5, no. 1, p. 3, 2018.
[23] P. Zhang, X. Xu, and H. Wang, "Computer-aided lung cancer diagnosis approaches based on deep learning," *J. Comput.-Aided Des. Comput. Graph.*, vol. 30, no. 1, p. 90, 2018.
[24] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imag.*, vol. 32, no. 4, pp. 582–596, 2019.
[25] H. Xiang, Y. Zhang, and H. U. Haibo, "Big data and security visualization," *J. Chongqing Univ.*, vol. 39, no. 2, pp. 71–81, 2016.
[26] J. G. Stadler, K. Donlon, J. D. Siewert, T. Franken, and N. E. Lewis, "Improving the efficiency and ease of healthcare analysis through use of data visualization dashboards," *Big Data*, vol. 4, no. 2, pp. 129–135, 2016.
[27] V. Satagopam, W. Gu, S. Eifes, P. Gawron, M. Ostaszewski, S. Gebel, A. Barbosa-Silva, R. Balling, and R. Schneider, "Integration and visualization of translational medicine data for better understanding of human diseases," *Big Data*, vol. 4, no. 2, pp. 97–108, 2016.
[28] A. Ledesma, M. Al-Musawi, and H. Nieminen, "Health figures: An open source JavaScript library for health data visualization," *BMC Med. Inform. Decis. Making*, vol. 16, no. 1, 2016, Art. no. 38.
[29] S. Siuly and Y. Zhang, "Medical big data: Neurological diseases diagnosis through medical data analysis," *Data Sci. Eng.*, vol. 1, no. 2, pp. 54–64, 2016.



YING YANG was born in Henan, China, in 1981. She received the bachelor's degree from the Xi'an University of Finance and Economics, in 2004, and the master's degree from the Wuhan University of Technology, in 2010. Since 2004, she has been with Xinxiang Medical University. She has published six articles. Her research interests include hospital engineering management and hospital operation management.

TAO CHEN was born in Henan, China, in 1981. He received the bachelor's degree from the Henan University of Technology, in 2005, and the master's degree from the Wuhan University of Technology, in 2010. Since 2005, he has been with The Third Affiliated Hospital of Xinxiang Medical University. He has published three articles. His research interests include hospital engineering management and hospital operation management.