Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

LANFANG SUN[1,2], XIN JIANG[1], HUIXIA REN[1,3], AND YI GUO[1]

[1] The Second Clinical Medical College (Shenzhen People's Hospital), Jinan University, Shenzhen 518000, China
[2] School of Software, Dalian University of Technology, Dalian 116000, China
[3] Clinical Medicine Postdoctoral Research Station, Jinan University, Guangzhou 510632, China

Corresponding authors: Xin Jiang (jiangxinsz@163.com), Huixia Ren (shengji054@163.com), and Yi Guo (xuanyi_guo@163.com)

ABSTRACT With the booming development of medical informatization and the ubiquitous connections in the fifth generation mobile communication technology (5G) era, the heterogeneity and explosive growth of medical data have brought huge challenges to data access, security and privacy, as well as information processing in Internet of Medical Things (IoMT). This article provides a comprehensive review of how to realize the timely processing and analysis of medical big data and the sinking of high-quality medical resources under the constraints of the existing medical environment and medical-related equipment. We mainly focus on the advantages brought by the cloud computing, edge computing and artificial intelligence technologies to the IoMT. We also explore how to rationalize the use of medical resources and the security and privacy of medical data, so that high-quality medical services can be provided to patients. Finally, we discuss the current challenges and possible future research directions in the edge-cloud computing and artificial intelligence related IoMT.

INDEX TERMS Internet of medical things (IoMT), deep learning, edge of computing, computation offloading.

I. INTRODUCTION

We know that although the digital transformation in the medical field started earlier, it has developed slowly compared with other fields [1]. With the rapid development of science, technology and economy, medical treatment has become one of the focuses of personal, social and even national attention. The traditional medical model has problems such as difficulty in seeing a doctor, expensive treatment, and occlusion of medical information. However, with the formal introduction of the concept of the Internet of Things (IoT) [2] in 1999, the application field of IoT has been involved in all aspects [3], [4] in today's era of Internet of Everything (IoE) [5]. As far as the medical field is concerned, Internet of Medical Things (IoMT) is the concentrated embodiment of IoT technology in this field, and also the core of medical digital transformation.

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

IoMT, that is, the complex IoT technologies such as radio frequency identification technology, sensor technology and positioning technology [6], are further applied to the medical field in combination with mobile terminals, network communication and other devices to realize the interaction between patients, medical staff, medical institutions and medical devices, to achieve the digitalization, automation and intelligence of the hospital.

The IoMT serves all aspects of the medical field, including identity recognition, vital signs monitoring, remote monitoring, medical drugs, waste and equipment monitoring. The wireless medical sensor is the foundation and core of each link, and the wireless sensor network [7] composed of many types of wireless medical sensors, such as pressure sensor, biosensor, implantable sensor, plays an important role in obtaining vital signs data of patients. Nowadays, sensors have been widely used in the operation room, emergency room and Intensive Care Unit (ICU) to monitor and display the

main conditions of patients. Besides, the wearable medical equipment with a sensor as the core [8] breaks through the limitation of time and space, provides real-time health monitoring for patients, greatly reduces the treatment cost of both doctors and patients, and enables patients to get personalized medical services anytime and anywhere.

We believe that the rise of the fifth generation mobile communication technology (5G) will usher in explosive development in the medical field, and the application scenarios of digital medicine such as remote consultation, operation guidance, first aid vehicle, wearable medical equipment will be more abundant [9], which also means that the sources of medical big data will be more diversified and grow rapidly. How to process, analyze and decide the collected medical data in real time and quickly is very important, which is directly related to the health and life of patients. In addition, since the medical data involves the personal privacy of patients, it is also important to strengthen the protection of sensitive data and patient privacy data.

Although the application of IoT technology in the medical field can help hospitals realize the intelligent medical treatment of people and the intelligent management of things, different medical institutions are relatively independent, so it is difficult to achieve resource sharing. The IoMT based on cloud computing provides powerful IT basic resources and greatly reduces medical costs. It can not only satisfy the mass storage of medical data, but also realize the sharing of medical information through the cloud platform, to improve the efficiency and quality of medical services. However, completely relying on cloud computing will consume huge network transmission resources and bring a huge delay, which is likely to pose a threat to the life of patients. The ability of cloud computing to process data sinks, making data processing closer to the source, rather than external data or cloud, which can shorten the delay time, and achieve real-time and faster processing and analysis of medical data. Edge computing reduces the dependence on a remote centralized server or distributed local server, and solves the problems existing in cloud computing through the reasonable application of resources on edge devices, which means that hospitals and clinics get a more agile and responsive IT network, so that patients can enjoy better medical services. But edge computing does not exist in isolation. Cloud computing and edge computing complement each other and play an extremely important role. Cloud computing focuses on the overall grasp, while edge computing focuses on the local. The rational use of edge cloud collaboration will better promote the development of medical application scenarios.

As far as we know, there are not many articles related to IoMT. Based on the existing environment in the medical field, this article combines three promising technologies for the first time to comprehensively review the existing research, and analyzes the challenges that the IoMT will face in the future. The main contributions of this article are as follows:

- First, we introduce the architecture of traditional IoMT, analyze the key IoT technologies applied in the medical field, and propose the research directions to maximize its effectiveness in the medical field. In addition, we also analyze the existing application research of the IoMT, and discuss the inevitable trend of the future development of the IoMT.
- Considering the rapid growth of medical data and the complexity of data structure, we study the three-tier architecture of cloud computing for IoMT and introduce the technologies involved in the application of cloud computing to IoMT. Also, we focus on the security and privacy of electronic health data in the cloud.
- By comparing with traditional IoMT and medical cloud IoT, we consider the advantages of edge computing and introduce its architecture in IoMT. Besides, the technology of computing offload which can provide faster and more efficient computing services in edge computing is discussed and the optimization direction is given.
- Finally, according to the limitations of current development, existing research content and emerging 5G technology, we discuss the possible future development directions and research challenges in the field of edge medical cloud

The main purpose of this article is to analyze how to deal with medical big data in real-time and quickly and how to make high-quality medical resources sink. The rest of this article is organized as follows: In Section II, we introduce the architecture and key technologies of the IoMT, and analyze how the IoT technology is applied to the medical field. In Section III, We mainly discuss the sharing of medical resources, medical big data processing and the security, privacy, and integrity of medical data based on cloud computing. In Section IV, We introduce the application of edge computing in the timely processing of medical data, the effective utilization of medical resources and the reduction of energy consumption. In Section V, We describe the current challenges faced by the edge medical cloud and the possible development direction in the future. Finally, we conclude this article in Section VI.

II. INTERNET OF MEDICAL THINGS

The application of IoT technology in the medical field realizes the digital management of medical information [10], so that medical staff is no longer busy with the recording and sorting of a large number of cumbersome medical information, but more focused on patients as the center to provide better medical services. In this section, we will introduce the architecture of the traditional IoMT, and then describe the key technologies used. Finally, we will introduce the research status of making the IoT better applied in the rapidly developing medical field.
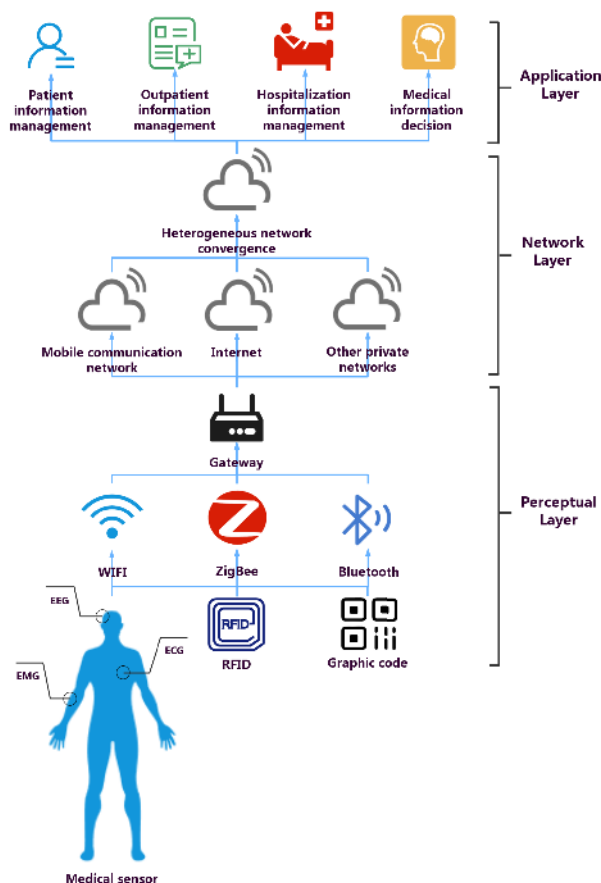
L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE*Access*

FIGURE 1. The architecture of IoMT.

A. ARCHITECTURE OF IoMT

At present, many scholars have studied the architecture of the IoT and put forward some architectures that can be applied to the IoT. There are mainly the following: web of things architecture, EPCglobal architecture, sensor network based architecture, autonomous architecture and Machine to Machine (M2M) architecture. Among them, M2M architecture covers some related contents of EPCglobal and Wireless Sensor Network (WSN) [11] and is the most widely used architecture in the field of IoT. The IoMT is the concentrated embodiment of IoT technology in the medical field. It follows the general three-tier architecture of IoT application, namely, perception layer, network layer, and transmission layer. Fig. 1 describes the architecture of the IoMT.

1) PERCEPTUAL LAYER

The perception layer is the focus and difficulty of the IoMT. It is mainly divided into two sublayers: data acquisition sublayer and data access sublayer. The sublayer of data acquisition is through different types of medical perception equipment and signal acquisition equipment to complete the perception and identification of nodes in the IoMT, as well as to collect the data information of people and things. It uses signal acquisition methods such as Radio Frequency Iden-

tification (RFID) technology, graphic code, image recognition technology, general packet radio service technology and multiple types of sensors such as physical signal sensor, physiological signal sensor, chemical sensor, DNA sensor to transmit all people and things involved in the network into Cyber-Physical Systems (CPS) nodes that are easy to identify. The nodes in the IoT are divided into three types: passive CPS, active CPS and Internet CPS. In the IoMT, corresponding identification is required according to different objects and needs. The data access sublayer uses a variety of access methods to connect the data collected by the data acquisition sublayer to the network layer through short-distance data transmission technology, such as ZigBee, Wi-Fi, Bluetooth, etc., and the main access methods should be selected according to the actual environmental characteristics of the IoMT and the needs of different objects.

2) NETWORK LAYER

The network layer is also divided into two sublayers: network transmission layer and service layer. The network transmission layer is the backbone network of the IoMT, which is equivalent to the nerve center and brain of human beings. It uses the mobile communication network, the Internet and other special networks to transmit the data information acquired by the perception layer in real-time, accurate, reliable and barrier free way. The formation of IoMT is not to completely replace the original heterogeneous networks, but to study the integration technology of heterogeneous networks suitable for hospitals and the original network. The service layer mainly realizes the integration of heterogeneous networks, the integration of various data formats, descriptions, data warehouses and other information. At the same time, it builds a service support platform based on this, which provides an open interface for various services in the application layer, so that the third party can develop relevant applications for the use of medical staff and other relevant personnel.

3) APPLICATION LAYER

The application layer is divided into two layers: medical information application and medical information decision-making application. Medical information application includes medical equipment and material information management, patient information management, outpatient information management, inpatient treatment information management, etc. The application of medical information decision-making includes patient information analysis, disease information analysis, medication information analysis, diagnosis and treatment information analysis, etc.

B. TECHNOLOGIES OF IoMT
1) RADIO FREQUENCY IDENTIFICATION (RFID)

RFID is a new non-contact automatic identification technology, which is one of the core technologies of the IoMT. It uses the radio frequency method to identify the designated target

IEEE Access

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

and read and write the relevant data, and the designated target and the identification system can complete the identification without physical contact, that is, non-contact two-way data communication. It has the advantages of long recognition distance, strong anti-interference and no need for human intervention. RFID system is generally composed of three parts: radio frequency electronic tag, reader and data management system. Radio frequency electronic tag is the core of RFID technology. Its main function is to store the relevant data information of the specified target and communicate with the reader and writer. RFID can achieve multi-target recognition and high-speed moving object recognition, which is widely used in medical asset management, medical equipment, waste tracking and personnel identification. It can also be used to collect vital signs data of patients, such as breath, blood pressure, ElectroCardioGram (ECG), etc., which is helpful for the relevant analysis of patients' condition. The low-cost ink-jet printed RFID tag antenna proposed in [12] can monitor the patient's medication dose, saving a lot of time and resources. In [13], considering the degradation of the reading range of the RFID tag when it is close to human tissue, a flexible RFID tag is proposed to expand its reading range. In real life, RFID tags will inevitably lose data and reduce recognition accuracy due to signal loss, tag breakage, channel conflict and other problems. A sparse representation classification algorithm based on dictionary segmentation in [14] can effectively solve this problem.

2) WIRELESS SENSOR NETWORK (WSN)

WSN integrates sensor technology [15], distributed information processing technology, communication technology and other technologies to achieve the three functions of data acquisition, processing and transmission. It is a wireless network system composed of a large number of small sensor nodes with wireless communication and computing capabilities, which are randomly deployed in or near the monitoring area through self-organization and multi-hop. It can monitor, perceive and collect all kinds of information of different environments or objects in real-time, and send the processed information to the client by wireless way. WSN has become a research hotspot in the medical field because of its reliability, rapidity, security, real-time and other advantages. Its application scope includes real-time monitoring of patients' physiological parameters, emergency monitoring, hospital general / ICU ward, etc. In order to improve the acquisition of patients' physiological parameters, a kind of graphene wearable medical sensor with high sensitivity and linearity is proposed in [16]. In [17], a new type of sensor is used, which can real-time monitor people's pressure level through their body temperature, movement speed and body perspiration, so as to reduce the risk of human health. In recent years, the optimization of WSN is also the focus of many researchers. In [18], aiming at the security problem of WSN, a low energy consumption network security mechanism based on the WSN smart grid monitoring application is proposed. To solve the energy consumption of WSN, Balanced-α Weighted Shortest Path (B-αWSP) routing algorithm is proposed in [19], which can reduce the energy consumption by effectively reducing the impact of high load transmission on node activities, while [20] is to minimize the energy consumption in the network through an optimal clustering algorithm based on compression sensing and principal component analysis. In a word, WSN technology has great development potential in the medical field.

3) MIDDLEWARE

Middleware plays a key role in the IoT [21], as well as in the IoMT. Middleware is located between the back-end application system and the reader-writer, which plays an intermediary role. It can meet the needs of a large number of applications, support distributed computing, and face sensor devices. By setting up a common interface and platform, it can realize the standardization of different application environments and the data communication between application systems. It will capture the data or events collected by the sensing device in real-time and conduct proofreading, filtering, collection and other processing, and then transmit them to the RFID reader or the back-end application database system to realize the data information interaction between the back-end application database system and the RFID reader. IoMT middleware adopts standard protocol and interface technology, which can develop different middleware for different medical application services and needs, such as electronic medical record information transmission middleware, medical staff management middleware, medical equipment management middleware, etc., and each middleware must be developed based on the requirements and standards of IoMT application services to achieve transmission standardization of data. Because of the large scale of the IoMT, high density of sensor nodes and complex business processes, the structure, transmission and processing methods of data collected by the sensing layer are very heterogeneous. In [22], we summarize the semantic middleware solution that combines technology and semantic technology to achieve a complete interoperability.

C. APPLICATIONS OF IoMT

The IoMT realizes the intelligent medical treatment and management of people and things, which not only reduces the cost of medical treatment, but also ensures people's health. Medical data has a high sensitivity. Whether the identity recognition system has enough security plays a crucial role in the IoMT platform. Single-mode biometric identification only has medium security and gradually can't control the explosive growth of data in the medical field. In [23], the multi-mode biometric identification which integrates the three biometrics of face, fingerprint and finger vein has a high recognition rate and higher security characteristics, which is also an inevitable trend in the future development of the medical field. In addition, vital signs monitoring based on [24] has brought great changes to personalized medical treatment. Real-time monitoring of vital signs [25]–[28] can

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE *Access*

timely understand the patient's physical condition [29] and analyze and make decisions on diagnosis and treatment information, which is helpful to study human diseases and provide preventive measures. In the case that patients do not have time to go or where they cannot visit doctors, the emerging telemedicine and mobile medicine based on smart-phones [30], laptops, tablet computers and other devices [31] have made great contributions, optimized the business process of the hospital, improved the work efficiency and resource utilization, which is the main direction of the application of the IoMT at present.

Whether it is real-time monitoring [32] of vital signs or the realization of telemedicine, it is inseparable from the support of wearable devices. The appearance of wearable devices such as rings, watches and wristbands makes the coordination between medical staff and patients more convenient. However, with the subversive changes in the medical field, the design speed, flexibility, and interconnection of wearable devices need to be further improved. A hybrid lean-agile method proposed in [33] is conducive to the development of wearable devices. In [34], the potential mechanism of human sweat leading to the separation of wearable devices is studied, which is helpful for the design of wearable devices applied to the skin, while the wearable large area touch sensor described in [35] can improve the adaptability and comfort of the wearer. With the rapid development of the medical field, the sense of the use of wearable devices will be valued.

III. CLOUD-ENABLED IoMT

The traditional IoMT is difficult to realize the interconnection between a large number of different medical institutions and medical information systems, which results in the isolated medical service information island, which will generate duplicate data in the medical process, lead to a large amount of waste of resources, and increase the interoperability difficulty between medical information systems. In addition, the data structure of massive medical data is various and complex, which puts forward higher requirements for data storage, processing, and management capabilities. Because medical data not only involves patients' privacy, but also affects disease control, informed decision-making and other aspects, the leakage of highly sensitive data will have a serious impact on individuals, families and even society. It is difficult for traditional IoMT to ensure the security, privacy, and integrity of massive data. In recent years, the application of cloud computing technology in the IoMT, with its high reliability, sharing, and scalability, has realized the leap forward reform in the medical field. In this section, we introduce the architecture and key technologies of the medical cloud IoT, and compare the security of the medical data in the cloud. Then, we introduce the latest research content of medical cloud IoT.

A. ARCHITECTURE OF CLOUD-ENABLED IoMT

Combined with the application and research of cloud computing, as shown in Fig. 2, the architecture of the medical cloud
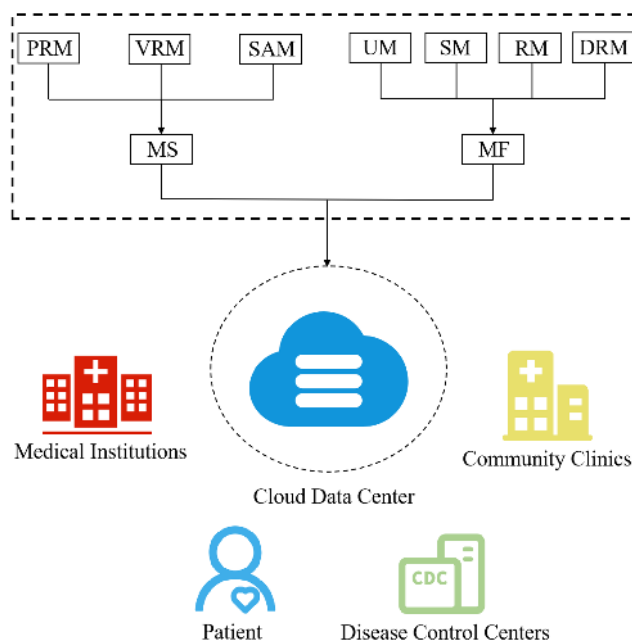

FIGURE 2. The architecture of the medical cloud IoT.

IoT is mainly divided into three layers: medical service layer, service management layer and user layer. The service management layer mainly includes a comprehensive Monitoring System (MS) for Physical Resources Monitoring (PRM), Virtual Resources Monitoring (VRM) and System Applications Monitoring (SAM), as well as Management Functions (MF) such as User Management (UM), Service Management (SM), Resource Management (RM), and Disaster Recovery Management (DRM).

1) MEDICAL SERVICE LAYER

It is mainly divided into three sublayers: medical infrastructure is a service layer, the medical platform is a service layer and medical software is a service layer. Medical infrastructure is the service layer, which provides powerful computing, storage, and network resources and services based on the hardware resource pool composed of hardware device virtualization. The construction of the digital hospital does not need to build its own data center, but can rent the required resources according to the needs, which will save a lot of costs for the large-scale hospital. Data center can be used to store personal data information of patients, statistical informa-tion of medical equipment and materials, medical analysis and decision-making information, etc., laying a foundation for data sharing, analysis and integration of medical institutions and systems. In addition, with the continuous growth of medical-related data and information, the required resources can also be expanded as a whole to achieve the flexible allocation of resources. For the user layer, the heterogeneity generated by the working mode of the layer and the construction process of the subsystem is invisible. A medical platform is the service layer, which provides a basic platform and corresponding

IEEE Access

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

technical support for the developers of the medical information system. Based on this platform, the developers can quickly build, develop and expand the functional modules of medical information system, which brings great convenience to the development, not only improves the development efficiency, but also saves the development cost. The medical software is the service layer, which realizes the release and interface of the functional software of the medical information system, mainly for the end-users who need to use the system. The functional software at this layer needs to provide requirements to the service management layer, and then form a more complete and accurate requirement for the platform as a service layer for development, and support the isolation of different user data and related configurations to ensure the security and privacy of medical data.

2) SERVICE MANAGEMENT LAYER

This layer is the core part of the medical cloud IoT and the foundation of the medical service layer. It ensures that the entire medical information system can operate safely and stably, and provides functions and technologies in maintenance, multiple management, and other aspects. This layer can play a good role in communicating with users, fully understand and analyze user needs, which is conducive to the medical platform as a service layer to avoid or eliminate early errors early, thereby improving the efficiency and quality of software development and reducing development costs. At the same time, this layer can also collect and screen all software in the medical platform as a service layer, and publish the screened software in the medical software as a service layer.

3) USER LAYER

Users at this level refer to users of medical information sharing systems, including medical institutions at different levels, disease control centers, community clinics, and so on. Users can access the user layer through the interface of mobile phones, computers, and other terminal devices, and will become a member of the cloud after access. While enjoying the services provided by cloud computing, users can also transmit their own resources to the cloud for other users to use, to realize real data sharing and resource sharing. For doctors, access to the cloud network can track and obtain patient information, treatment plans, treatment details, etc. in real-time, while patients can choose any medical institution for treatment according to their own conditions through the cloud network. More importantly, experts from different medical institutions can conduct collaborative consultation for severe patients or difficult and miscellaneous diseases through a cloud network, which greatly improves the possibility of curing patients and solving medical problems.

B. TECHNOLOGIES OF CLOUD-ENABLED IoMT
1) CLOUD COMPUTING

With the rapid development of modern science and technology, the advent of big data and information age, the infor-

mation generated in the medical field is becoming more and more complex, the heterogeneity of medical data is also increasing, and the access of data, information processing, and system maintenance are facing great challenges. This makes it particularly important to apply information management technology to the daily management and service process of medical institutions. In our daily life, we can use water, electricity, gas, etc. at a lower cost and unlimited, as well as cloud computing. According to the scope of application, cloud computing can be divided into a private cloud, public cloud, and hybrid cloud. According to the service type, cloud computing can be divided into infrastructure as a service, platform as a service and software as a service. It is a kind of distributed computing. Based on the virtual-ization technology, a large number of computing resources are collected to form a resource pool, which is managed and scheduled uniformly. These resources include computing resources, storage resources, network resources, etc. Compared with the traditional network application mode, cloud computing has the advantages of service resource pooling, resource sharing, high reliability, and high-cost performance. The resources in the cloud are infinitely scalable. Users do not need to understand the internal structure of the cloud. They only need to choose the corresponding resources according to their own needs and pay on demand, so they can get unlimited resources at any time and space. The rise and maturity of cloud computing technology provides a good opportunity for the realization of the digital hospital, and also brings great changes to the calculation, processing, and sharing of medical information and data. Unfortunately, cloud computing technology brings convenience, but at the same time, it is difficult for users to know whether the cloud service is safe or not, so they can not effectively control the whole. In [36], a quantifiable security evaluation model is proposed to solve this problem. The security of the data stored in the cloud server is very important. In [37], the dynamic audit scheme of big data using a fuzzy identity ensures the integrity of user data.

2) BIG DATA

With the advent of the cloud era, data has far exceeded the capabilities of traditional database software tools, and has the characteristics of authenticity, massive data scale, fast speed, diverse data types, and low-value density. Therefore, it is necessary to develop a new technology and management system that can not only control the huge data information, but also deal with the meaningful data professionally, so the big data technology came into being. In the field of medical treatment, a large number of patient information data, clinical experiment data, medical environment data and so on form the medical big data. Big data technology generally includes data collection and preprocessing, data storage, data analysis, and mining. Traditional data collection sources are single, and the amount of data stored, managed and analyzed is relatively small. Most of them can be processed through relational database and parallel database, while medical big data collection is generally divided into [38]. Data types

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE *Access*

are divided into structured, semi-structured and unstructured, and the size is often in TB or PB. The collected data is analyzed, extracted, cleaned and other operations to obtain high-quality data. Data storage needs to store the collected data to the corresponding database for management and call. The storage mode is usually divided into structured storage, semi-structured storage, and unstructured storage. Different storage modes are implemented in different ways. Structured storage is generally stored by MySQL, Oracle, SQL Server and other relational databases, semi-structured storage is implemented by the hadoop framework, and unstructured storage is implemented by NoSQL technol-ogy. Data analysis [39] is different from the traditional online analysis and processing, and its technology include [40], focusing on the breakthrough in visual analysis, data mining algorithm, predictive analysis, semantic engine, data quality, and data management. However, data mining is a process of extracting hidden but potentially valuable information and knowledge from a large number of incomplete, noisy, fuzzy and random data. At present, the decision tree, support vector machine and artificial neural network are the most popular data mining technologies. Using big data technology to effectively mine and research [41] medical big data will make a great contribution to disease diagnosis and treatment and medical research and become the source of promoting the construction of digital hospitals.

3) ARTIFICIAL INTELLIGENCE (AI)

AI [42] is a new technology science that can simulate and expand the theory, method, technology, and application system of human's thinking process and intelligent behavior [43]. It is generally divided into weak AI, strong AI, and super AI. The key technologies of AI include Machine Learning (ML) [44], [45], Natural Language Processing (NLP), robot, computer vision and expert system. Among them, ML is the most core technology in AI. It discovers potential laws through learning from a large number of data to guide human decision-making [46]–[48]. However, Deep Learning (DL) [49], [50] based on neural networks is the most representative branch of ML technology. Its structure is mostly multi-layer perceptron with multiple hidden layers, which is easy to find the deeper rules hidden in the data and has strong feature extraction ability [51]. Typical deep learning [52] models include Convolutional Neural Network (CNN), deep-belief net and stacked autoencoder network. In the medical field, ML technology can predict and diagnose diseases, which largely avoids the high error, low efficiency and the emergence of major diseases of artificial diagnosis. In [53], an emotion recognition system based on electroencephalo-gram (EEG), ECG and other medical sensors and CNN are proposed, which is helpful to the treatment of mental diseases. An expert system is an AI system including image recognition, a large number of knowledge and experience in specific fields in the medical field. It is also a representative application of AI technology in the medical field, with inspiration, pertinence, transparency, and flexibility. It can

simulate the decision-making process of mental activity and thinking activity when medical experts diagnose diseases, provide reliable and valuable medical diagnosis assistance, and greatly improve the diagnosis and treatment efficiency of patients. In addition, NLP can transform a large number of unstructured medical text information into structured data containing important medical information. The application of medical robots greatly improves the effect of diagnosis and treatment and nursing. Computer vision can quickly read and diagnose the medical image data such as microscope image, X-ray image, and UltraSound (US) image. The application of AI technology in the field of medical treatment [54], not only brings technological innovation, but also brings the change of medical service mode.

C. APPLICATIONS OF CLOUD-ENABLED IoMT

The construction of medical information based on cloud computing greatly alleviates the shortage of medical resources, reduces the cost of patients' medical treatment, and improves the level and efficiency of medical services. The core of the medical cloud IoT is to establish a medical information sharing platform based on the electronic health records, so as to solve the problem of information island in different medical institutions. It is of great value for hospitals and patients to upload electronic health files formed by personal information, vital signs and other relevant data of patients to the cloud for storage and unified management. However, whether the cloud electronic health data is absolutely safe and private [55] is the focus of medical scholars. In [56], a mobile medical application based on IBBE algorithm is proposed, which can encrypt the patient's health data before it is stored in the cloud, and the doctor must decrypt it through the identity key to obtain the patient's relevant health data. In [57], the mobile medical cloud architecture based on medibchain encrypts the patient's health data through two dynamic anonymities and one static anonymity. The low-cost authentication mechanism is realized by the shared key, and has high robustness. However, the LPP-MSA scheme proposed in [58] uses online / offline signature and server assisted verification mechanism to realize the privacy protection of patients with low computing and storage costs. In addition, telemedicine [59] is also an important application direction of cloud computing in the medical field. Through the pay as you go service-model, the cost of purchasing and maintaining the hardware needed to build telemedicine is greatly reduced, and the efficiency of telemedicine is improved. At the same time, it overcomes the distance barrier so that patients in remote areas or relatively poor areas can receive expert consultation by sending images or real-time video and carry out treatment and nursing under their guidance, to improve the level and quality of patients' medical treatment and even save their lives. Similarly, telemedicine also involves the safety of patient health data [60]. In [61], the KP-ABE scheme is proposed. It is a dynamic authentication and authorization scheme based on attribute encryption algorithm and blockchain technology [62] and applied to a

TABLE 1. The comparison of the security of cloud medical data.

Factors	Cloud Computing	Edge Computing
Computing architecture	Centralized	Distributed
Server node location	Edge network	Data center
Transmission bandwidth load	High	Low
Energy consumption	High	Low
Data processing	Slow	Fast
Latency	High	Low
Real time	Weak	Strong
Security	Low	High
Reliability	High	Low
Computing resources	Unlimited	Limited
Computing cost	High	Low
User experience	Weak	Strong

distributed telemedicine system. The scheme has high flexibility, fast data processing speed and can effectively protect the safety and integrity of patient health data. The application of medical cloud IoT can effectively realize patient-centered medical services, and ensuring the security of cloud data is the focus of future research. In TABLE 1., we compare the security of cloud medical data.

IV. EDGE-CLOUD IoMT

The IoMT based on cloud computing technology uploads the medical data from the terminal equipment to the remote cloud, and returns the results to the terminal equipment after calculation. However, if the huge amount of data generated by the growing medical equipment is uploaded to the cloud computing, it will cause huge pressure on the cloud, resulting in high energy consumption and huge delay due to the high load of the cloud. Cloud computing alone cannot help with such a large data set and provide real-time response. Therefore, the key to the development of medical cloud IoT is to effectively expand the ability of cloud computing and use distributed computing resources, that is, to process computing tasks at the edge of the network [69], and the application of edge computing can just meet this computing demand [70]. In this section, we introduce the architecture of the edge medical cloud IoT and its key technologies, and compare cloud computing with edge computing. Finally, we discuss the latest research content.

A. ARCHITECTURE OF EDGE-CLOUD IoMT

We divide the edge cloud IoMT into a terminal layer, edge computing layer and cloud computing layer. Fig. 3 describes the architecture of the edge medical cloud IoT.

The terminal layer consists of a variety of medical-related IoT devices, such as medical sensors, wearable devices, RFID tags, etc. It is the closest layer to the end-user, mainly responsible for collecting data from the local device and uploading the data to the edge device with the input mode as the carrier. The edge computing layer consists of a large number of network edge nodes, which can be intelligent terminal devices, such as smart-phones, tablets, etc., or network
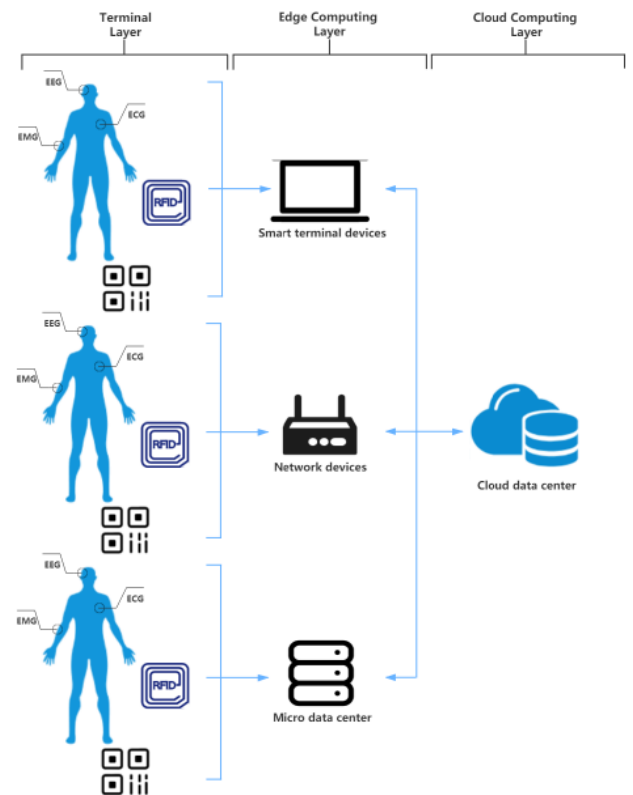
FIGURE 3. The architecture of the edge medical cloud IoT.

devices, such as gateways, routers, etc. These edge nodes are widely deployed between the terminal equipment and the cloud, such as hospitals, clinics, etc., they can provide edge computing, storage and network services for the received data, and because of the small number of hops between the edge nodes and the terminal equipment, hospitals, clinics, etc. can obtain a more agile and responsive network. This can reduce request delay, effectively avoid data leakage caused by long-distance transmission and other security issues, which is critical for extremely sensitive medical data. In addition, the layer regularly sends the processed data to the cloud for subsequent analysis. The cloud computing layer is in the center of the whole network, and it is a powerful data processing center, with massive computing, storage, and network resources, which can summarize, analyze and permanently store the data uploaded by the edge computing layer. At the same time, the deployment strategy of edge computing layer can be dynamically adjusted and distributed according to the network resources, which can improve the quality of delay-sensitive services of edge nodes. Fig. 4 describes the models of cloud computing and edge computing.

B. TECHNOLOGIES OF EDGE-CLOUD IoMT
1) EDGE COMPUTING

Edge computing [72] refers to an open platform that provides the nearest service with the core capabilities of a network, computing, storage and application on the edge of the
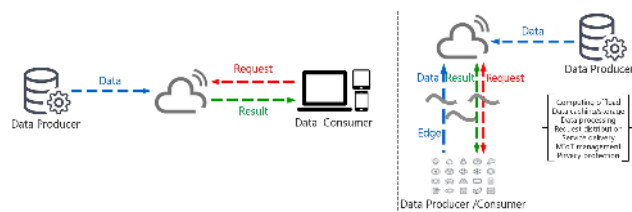
L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE *Access*

FIGURE 4. The models of cloud computing and edge computing [71].

network close to the source of people, objects or data. That is to say, the computing and storage capacity of the cloud will sink to the edge of the network, and the original data will not be uploaded to the cloud for centralized processing as much as possible, but the analysis and processing of data sources will be realized locally through distributed computing and storage. And the difference between edge computing and cloud computing lies in the two-way communication between terminal equipment and cloud computing center, that is, terminal equipment can send requests to cloud computing center, can also accept the unified control, processing and data storage of cloud computing center, and complete the computing tasks and adjustment strategies issued by cloud computing center. Edge computing technology has the characteristics of low delay, high security, definable, schedulable and so on. When the network is abnormal or even directly interrupted, edge nodes can realize local autonomy and self-recovery and have high robustness. The calculator system proposed in [73] can make full and effective use of the computing resources of edge nodes to obtain valuable data information in real time. In addition, edge computing can be divided into Moving Edge Computing (MEC) [74], [75], Micro Cloud Computing (MCC) and fog computing. Among them, MEC is also called multi-access edge computing. It uses the edge of mobile networks to provide services and cloud computing functions for users nearby, but the limited computing power and maintenance cost are the issues MEC needs to pay attention to. In [76], it is mentioned that the unloading scheme is very important for balancing task demand and budget, and an equilibrium pricing strategy based on the approximate greedy algorithm is proposed for the unloading scheme, which can maximize task utility [77]. MCC is like a small mobile data center, which is deployed on the edge of the network, and can provide real-time resources for users, while fog computing [78], [79] is composed of computing devices with weak performance and dispersion. It adopts a distributed and closer to the edge of the network architecture, and can basically provide all kinds of applications that cloud computing can provide, but its computing power is weak. These three forms are similar in deployment location, application scenario, real-time interaction, etc., but they are different [80].

2) COMPUTATION OFFLOADING
Due to the limitations of computing power, storage space and battery life of mobile devices, it can not meet the

operation requirements of a large number of emerging mobile applications such as low latency, high efficiency, and high reliability. In mobile cloud computing, computing offload [81] technology is proposed for the first time. It offloads some or all computing tasks of heavy load and computing-intensive mobile devices to the cloud server, and uses its powerful computing power and rich storage resources to process computing tasks and return the results to users to enhance the data processing ability of mobile devices and ease the storage of resources limit and reduce its energy consumption. However, there are some problems such as unpredictable delays, data security, long-distance transmission energy consumption, etc. Compared with offloading computing tasks to cloud servers, computing offloading in edge computing can provide faster and more efficient computing services for terminal devices by offloading computing tasks to network edge servers [82]. Computational offload technology mainly includes two problems: decision-making and resource allocation. Due to the availability of cloud servers, the performance of mobile devices and the usage habits of users, it is very important to make a proper decision of offload. In [83], the decision of computing offload based on game theory and the sub-optimal algorithm based on Game Theory in [84] can be used in multiple access edge computing to save the overall computing cost and achieve the Nash equilibrium through a limited step size. However, in [85], the problem of minimizing the energy consumption of edge computing based on non-orthogonal multiple access technologies is considered. After the decision-making of unloading, the next consideration is the rationality of resource allocation. Resource allocation mainly studies whether to unload the computing task to one or more MEC servers, which depends on whether the computing task can be divided and whether there is a correlation between the divided parts. In [86], a resource allocation method based on core MEC servers is proposed, which can solve the problem of transmission delay caused by resource allocation among multiple MEC servers. The resource allocation scheme based on the deterministic differential equations-in [87] can effectively allocate resources and ensure the security and stability of the MEC application environment.

C. APPLICATIONS OF EDGE-CLOUD IoMT
In the field of medical treatment, the rapid, efficient and reliable IT infrastructure determines the health and life of patients. When doctors use medical equipment, they need to obtain the basic information and vital sign data of patients in real-time and quickly to make the most beneficial medical decisions for patients, which may be life and death related. However, the existing cloud infrastructure has been unable to effectively manage the exponential growth of patient health data generated by medical devices. Once there is a network outage and it cannot be solved for the first time, it will have a great impact on the normal work of the hospital, and even affect the lives of patients. Edge computing, in the form of a decentralized distributed system, improves the working

IEEE *Access*

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

TABLE 2. The comparison of cloud computing and edge computing.

Factors	Cloud Computing	Edge Computing
Computing architecture	Centralized	Distributed
Server node location	Edge network	Data center
Transmission bandwidth load	High	Low
Energy consumption	High	Low
Data processing	Slow	Fast
Latency	High	Low
Real time	Weak	Strong
Security	Low	High
Reliability	High	Low
Computing resources	Unlimited	Limited
Computing cost	High	Low
User experience	Weak	Strong

efficiency of medical staff and the medical experience of patients. Medical staff does not need to send the patient health data to the remote data center for processing and waiting for the return results. They can set up the edge computing data center to process, calculate and store the patient health data collected by the edge devices locally, which brings great changes to the medical field. Edge computing enables wearable health monitoring equipment to perform real-time analysis and processing locally even when it is offline, which is helpful for remote patient monitoring [88] and inpatient care. In [89], a health monitoring architecture based on edge computing, AI and other technologies is proposed. This architecture can be used to collect ECG, blood pressure, EEG and other types of vital signs data and improve the accuracy of disease diagnosis. However, in [6], a distributed method with scalable clustering technology is used. This method can realize the privacy and accuracy analysis of medical data by aggregating and analyzing a large number of heterogeneous medical data collected in edge devices before sending them to the cloud. The health system based on edge computing in [90] can reduce system delay, reduce energy consumption, and optimize the transmission of medical data. Besides, for medical Augmented Reality (AR) / Virtual Reality (VR) and other businesses, edge computing can quickly process patient interaction information and surgical image materials, providing doctors with visual assistance and higher surgical accuracy and success rate. In [91], an MCC platform based on virtual sensors is proposed for the collection and analysis of patient health data. Compared with physical sensors, virtual sensors can solve the problems of heterogeneous physical sensors, resource and processing limitations. The health monitoring system based on fog computing in [92] not only reduces the data flow generated by the network core, but also ensures the safety of patient health data stored in the local area.

TABLE 2. shows the comparison between cloud computing and edge computing. We say that edge computing is actually an extension of the concept of cloud computing, which can not completely replace cloud computing. The relationship between edge computing and cloud computing is collaborative and complementary. The edge ends can analyze and process a large number of real-time data quickly, but most of the data is not only used once. Even after the edge end processing, it still needs to be collected from the edge end to the cloud. The mining and analysis of massive data, the storage of key data and the linkage of multiple edge nodes all need to rely on the cloud, and the virtualization resources and management of the edge also need to be completed by the cloud. Only when edge computing and cloud computing work closely together can they achieve different demand scenarios, thus maximizing the application value of edge computing and cloud computing.

V. RESEARCH CHALLENGE AND FUTURE DIRECTION FOR EDGE-CLOUD COMPUTING ENABLED IoMT
A. EDGE MEDICAL CLOUD BASED ON 5G
Compared with 4G, 5G has the characteristics of high speed, low delay, wide connection, faster mobile speed, higher security and more flexible service deployment, which has brought great impact on the innovation of edge computing technology. In addition, the combination of 5G and edge computing has laid a technical foundation for the development of smart medicine, especially telemedicine. Telemedicine, including remote consultation, remote surgery, and remote US, requires real-time ultra-high-definition image quality, medical image and other massive data-transmissions, which also promotes the deployment and implementation of "5G + edge computing". At present, there have been successful cases of cross-domain remote precise operation control and guidance. "5G + edge com-puting" is of great significance to the subsidence of high-quality medical resources, the alleviation of uneven distribution of medical resources, and the reduction of patients' medical costs. With the arrival of the 5G era and the continuous development of edge computing, the future of telemedicine means will emerge in endlessly. However, 5G also brings a lot of security problems, including security capacity opening, virtualization security, heterogeneous paradigm authentication and authentication, and other native challenges, VR / AR content security, traditional security, data security and privacy protection and other application challenges, which are all worthy of study in the future.

B. EDGE DEVICE SECURITY AND PRIVACY
We know that due to the performance, cost, resource limitations and scattered geographical location of edge devices, they are more vulnerable to data tampering, privacy disclosure, malicious nodes, and other service facilities attacks, denial of service, middleman, malicious gateway and other network facilities attacks, as well as virtualization facilities attacks and many other security and privacy threats. In order to avoid or reduce the occurrence of such problems, the improvement and extension of hash function, symmetric and asymmetric encryption algorithm and attribute encryption algorithm are very important. For the authentication of

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE *Access*

edge nodes, also to a single edge node, the mutual authentication of multiple nodes should be considered. In addition, the development of communication security protocol, the trust management of edge nodes and the distributed edge intrusion detection technology to prevent malicious attacks are also the focus of future research. For the security and privacy problems that have occurred, it is also very important to develop the corresponding edge node isolation mechanism or the scheme that can make the service facilities continue to serve.

C. EDGE CACHING AND ENERGY CONSUMPTION

Because the caching of edge nodes is done locally, it can effectively relieve the pressure of the core network and has low network communication overhead, interaction delay and bandwidth cost. However, the cache capacity of edge nodes is limited. How to determine the cached content, cache strategy and cache update strategy is an urgent problem. Making full use of multiple edge nodes for collaborative caching and adaptive caching is a promising solution, but it also brings challenges to the establishment of a trust management framework. How to select reliable and secure edge nodes and the credibility update strategy of edge nodes are the key to improve the quality of the cache. It is one of the current research directions to integrate 5G technology and blockchain technology and apply them to edge cache. The application of ML and big data technology also provides feasibility for the adjustment of the cache strategy, and the cache strategy for video content is the difficulty and hotspot. In addition, due to the limitation of edge devices, how to reduce energy consumption is also a problem to be solved. In addition to the cooperative caching of edge nodes that can effectively reduce energy consumption, we also need to consider the strategy of computing offload and how to reduce the computing overhead of edge devices. The higher the computational complexity is, the greater the energy consumption will be. Therefore, the design of an effective and low complexity algorithm is very important to reduce energy consumption.

D. OPTIMIZATION OF AI

The development of edge computing benefits from the promotion of AI. Most of the edge of intelligent computing depends on the AI chip integrated on the edge device. The application of AI algorithms in edge computing makes it faster, safer and more efficient to process data, and it can effectively allocate edge resources, to minimize the cost of edge services. At present, DL is the main research direction of AI, which needs to focus on how to optimize the algorithm by solving its non convex optimization problem, gradient vanishing problem and over fitting problem. In fact, most of the objective functions in DL are complex, so there are many optimization problems without analytic solutions. Therefore, it is a good solution to find approximate solutions using the optimization algorithm based on numerical methods such as random gradient descent. However, how to escape from local optimum in low dimensional space and saddle point in high

dimensional space are two great challenges in optimization. Only by continuously optimizing the algorithm, can the edge devices achieve the maximum utility, and then better serve various application scenarios in the medical field.

VI. CONCLUSION

In this article, we analyzed the traditional IoMT, the cloud-based IoMT and the edge-based IoMT, focusing on the medical data processing of edge medical cloud and the development of telemedicine. Specifically, we first introduced the architecture of traditional IoMT and the key technologies involved, discussed the application of IoT technology in the medical field and the problems involved, and gave the optimization direction. On this basis, we discussed the problems of traditional IoMT and the advantages of applying cloud computing to IoMT, analyzed the key technologies of cloud computing, and focused on the data security of medical cloud. Then, we analyzed the advantages of edge computing compared with cloud computing and discussed the optimization of edge computing, and proposed that edge cloud collaboration can achieve maximum utility in the medical field. We believe that with the continuous development of the medical field, to ensure the security of medical data and the privacy of patients, as well as how to reduce energy consumption, is an urgent problem to be solved, and also a major challenge to be faced in the future.

REFERENCES

[1] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: Enabling smart mobility," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 16–55, May 2017.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[3] W. Wang, S. Yu, T. M. Bekele, X. Kong, and F. Xia, "Scientific collaboration patterns vary with scholars' academic ages," *Scientometrics*, vol. 112, no. 1, pp. 329–343, Jul. 2017.

[4] X. Wang, Z. Ning, X. Hu, L. Wang, L. Guo, B. Hu, and X. Wu, "Future communications and energy management in the Internet of Vehicles: Toward intelligent energy-harvesting," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 87–93, Dec. 2019.

[5] Z. Ning, X. Hu, Z. Chen, M. Zhou, B. Hu, J. Cheng, and M. S. Obaidat, "A cooperative quality-aware service access system for social Internet of vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2506–2517, Aug. 2018.

[6] R. C. Shit, S. Sharma, D. Puthal, and A. Y. Zomaya, "Location of things (LoT): A review and taxonomy of sensors localization in IoT infrastructure," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2028–2061, 3rd Quart., 2018.

[7] X. Liu, A. Liu, T. Qiu, B. Dai, T. Wang, and L. Yang, "Restoring connectivity of damaged sensor networks for long-term survival in hostile environments," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1205–1215, Feb. 2020.

[8] G. Yang, G. Pang, Z. Pang, Y. Gu, M. Mäntysalo, and H. Yang, "Non-invasive flexible and stretchable wearable sensors with nano-based enhancement for chronic disease care," *IEEE Rev. Biomed. Eng.*, vol. 12, pp. 34–71, 2019.

[9] J. Yu, J. Liu, R. Zhang, L. Chen, W. Gong, and S. Zhang, "Multi-seed group labeling in RFID systems," *IEEE Trans. Mobile Comput.*, early access, Aug. 14, 2019, doi: 10.1109/TMC.2019.2934445.

[10] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, and R. Y. K. Kwok, "Mobile edge computing enabled 5g health monitoring for Internet of medical things: A decentralized game theoretic approach," *IEEE J. Sel. Areas Commun.*, pp. 1–16, 3rd Quart., 2020.

[11] X. Liu, T. Qiu, and T. Wang, "Load-balanced data dissemination for wireless sensor networks: A nature-inspired approach," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9256–9265, Dec. 2019.

[12] A. Sharif, J. Ouyang, Y. Yan, A. Raza, M. A. Imran, and Q. H. Abbasi, "Low-cost inkjet-printed RFID tag antenna design for remote healthcare applications," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 3, no. 4, pp. 261–268, Dec. 2019.

[13] S. M. A. Morais, A. J. R. Serres, J. I. L. Araujo, C. C. R. de Albuquerque, C. C. P. e Silva, M. L. Ferreira, G. K. F. Serres, H. A. Pereira, and A. M. de Oliveira, "Flexible RFID tag with circular CSRR for body-area applications," in *Proc. 4th Int. Symp. Instrum. Syst., Circuits Transducers (INSCIT)*, Aug. 2019, pp. 1–4.

[14] X. Luo, Q. Yang, P. Li, T. Miyazaki, and X. Wang, "Grouping strategy for RFID-based activity recognition in smart home," in *Proc. Int. Conf. Internet Things (iThings), IEEE Green Comput. Commun. (GreenCom), IEEE Cyber, Phys. Social Comput. (CPSCom), IEEE Smart Data (SmartData)*, Jul. 2019, pp. 96–101.

[15] W. Wei and Y. Qi, "Information potential fields navigation in wireless ad-hoc sensor networks," *Sensors*, vol. 11, no. 5, pp. 4794–4807, May 2011.

[16] Y. Pang, Z. Yang, Y. Yang, X. Wu, Y. Yang, and T.-L. Ren, "Graphene based wearable sensors for healthcare," in *Proc. Int. Conf. IC Design Technol. (ICICDT)*, Jun. 2019, pp. 1–4.

[17] L. Rachakonda, P. Sundaravadivel, S. P. Mohanty, E. Kougianos, and M. Ganapathiraju, "A smart sensor in the IoMT for stress level detection," in *Proc. IEEE Int. Symp. Smart Electron. Syst. (iSES) (Formerly iNiS)*, Dec. 2018, pp. 141–145.

[18] G. S. Dhunna and I. Al-Anbagi, "A low power WSNs attack detection and isolation mechanism for critical smart grid applications," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5315–5324, Jul. 2019.

[19] N. Bouadem, R. Kacimi, and A. Tari, "B-d'WSP selection algorithm: A load balancing convergecast for WSNs," in *Proc. Wireless Days (WD)*, Apr. 2018, pp. 101–103.

[20] S. Yadav, V. Kumar, S. B. Dhok, and D. N. K. Jayakody, "Energy-efficient design of MI communication-based 3-D non-conventional WSNs," *IEEE Syst. J.*, early access, Jun. 11, 2019, doi: 10.1109/JSYST.2019.2918184.

[21] M. A. A. da Cruz, J. J. P. C. Rodrigues, J. Al-Muhtadi, V. V. Korotaev, and V. H. C. de Albuquerque, "A reference model for Internet of Things middleware," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 871–883, Apr. 2018.

[22] R. Zgheib, E. Conchon, and R. Bastide, "Semantic middleware architectures for iot healthcare applications," in *Enhanced Living Environments*. New York, NY, USA: Springer, 2019, pp. 263–294.

[23] Y. Xin, L. Kong, Z. Liu, C. Wang, H. Zhu, M. Gao, C. Zhao, and X. Xu, "Multimodal feature-level fusion for biometrics identification system on iomt platform," *IEEE Access*, vol. 6, pp. 21418–21426, 2018.

[24] S. Banou, M. Swaminathan, G. Reus Muns, D. Duong, F. Kulsoom, P. Savazzi, A. Vizziello, and K. R. Chowdhury, "Beamforming galvanic coupling signals for IoMT Implant-to-Relay communication," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8487–8501, Oct. 2019.

[25] Z. Ning, K. Zhang, X. Wang, L. Guo, X. Hu, J. Huang, B. Hu, and R. Y. K. Kwok, "Intelligent edge computing in internet of vehicles: A joint computation offloading and caching solution," *IEEE Trans. Intell. Transport. Syst.*, pp. 1–14, May 2020, doi: 10.1109/TITS.2020.2997832.

[26] F. Lamonaca, D. L. Carni, V. Spagnuolo, G. Grimaldi, F. Bonavolontà, A. Liccardo, R. S. L. Moriello, and A. Colaprico, "A new measurement system to boost the IoMT for the blood pressure monitoring," in *Proc. IEEE Int. Symp. Meas. Netw. (M N)*, Jul. 2019, pp. 1–6.

[27] F. Lamonaca, E. Balestrieri, I. Tudosa, F. Picariello, D. L. Carnì, C. Scuro, F. Bonavolontà, V. Spagnuolo, G. Grimaldi, and A. Colaprico, "An overview on Internet of medical things in blood pressure monitoring," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2019, pp. 1–6.

[28] W. Young, J. Corbett, M. S. Gerber, S. Patek, and L. Feng, "DAMON: A data authenticity monitoring system for diabetes management," in *Proc. IEEE/ACM 3rd Int. Conf. Internet-of-Things Design Implement. (IoTDI)*, Apr. 2018, pp. 25–36.

[29] P. Kumar Das, F. Zhu, S. Chen, C. Luo, P. Ranjan, and G. Xiong, "Smart medical healthcare of Internet of Medical Things (IOMT): Application of non-contact sensing," in *Proc. 14th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2019, pp. 375–380.

[30] Z. Ning, F. Xia, X. Hu, Z. Chen, and M. S. Obaidat, "Social-oriented adaptive transmission in opportunistic Internet of smartphones," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 810–820, Apr. 2017.

[31] X. Huang, Y. Li, J. Chen, J. Liu, R. Wang, X. Xu, J. Yao, and J. Guo, "Smartphone-based blood lipid data acquisition for cardiovascular disease management in Internet of Medical Things," *IEEE Access*, vol. 7, pp. 75276–75283, 2019.

[32] Y. Sun, H. Qiang, J. Xu, and G. Lin, "Internet of Things-based online condition monitor and improved adaptive fuzzy control for a medium-low-speed maglev train system," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2629–2639, Apr. 2020.

[33] N. Glazkova, C. Fortin, and T. Podladchikova, "Application of lean-agile approach for medical wearable device development," in *Proc. 14th Annu. Conf. Syst. Syst. Eng. (SoSE)*, May 2019, pp. 75–80.

[34] Q. Wang, W. Chen, and J. Wu, "Effect of capillary bridges on the interfacial adhesion of wearable electronics to epidermis," *Int. J. Solids Struct.*, vols. 174–175, pp. 85–97, Nov. 2019.

[35] P. S. Pessoa, R. C. Giacomini, C. Borelli, and R. T. Buhler, "Wearable large-area touch sensor: Design, simulation and experimental results," in *Proc. 4th Int. Symp. Instrum. Syst., Circuits Transducers (INSCIT)*, Aug. 2019, pp. 1–5.

[36] A. Sun, G. Gao, T. Ji, and X. Tu, "One quantifiable security evaluation model for cloud computing platform," in *Proc. 6th Int. Conf. Adv. Cloud Big Data (CBD)*, Aug. 2018, pp. 197–201.

[37] C. Zhao, L. Xu, J. Li, F. Wang, and H. Fang, "Fuzzy identity-based dynamic auditing of big data on cloud storage," *IEEE Access*, vol. 7, pp. 160459–160471, 2019.

[38] M. Li, C. Wang, L. Yan, and S. Wei, "Research on the application of medical big data," in *Proc. 14th Int. Conf. Comput. Sci. Edu. (ICCSE)*, Aug. 2019, pp. 478–482.

[39] H. H. G. Noriega and F. García-Sánchez, "Semantic (Big) data analysis: An extensive literature review," *IEEE Latin Amer. Trans.*, vol. 17, no. 5, pp. 796–806, May 2019.

[40] F. Amalina, I. A. Targio Hashem, Z. H. Azizul, A. T. Fong, A. Firdaus, M. Imran, and N. B. Anuar, "Blending big data analytics: Review on challenges and a recent study," *IEEE Access*, vol. 8, pp. 3629–3645, 2020.

[41] Y. Yu, M. Li, L. Liu, Y. Li, and J. Wang, "Clinical big data and deep learning: Applications, challenges, and future outlooks," *Big Data Mining Analytics*, vol. 2, no. 4, pp. 288–305, Dec. 2019.

[42] K. Zhang, S. Leng, X. Peng, L. Pan, S. Maharjan, and Y. Zhang, "Artificial intelligence inspired transmission scheduling in cognitive vehicular communications and networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1987–1997, Apr. 2019.

[43] W. Qi, B. Landfeldt, Q. Song, L. Guo, and A. Jamalipour, "Traffic differentiated clustering routing in DSRC and C-V2X hybrid vehicular networks," *IEEE Trans. Veh. Technol.*, early access, Apr. 28, 2020, doi: 10.1109/TVT.2020.2990174.

[44] W. Wang, J. Chen, J. Wang, J. Chen, and Z. Gong, "Geography-aware inductive matrix completion for personalized point-of-interest recommendation in smart cities," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4361–4370, May 2020.

[45] W. Wang, J. Chen, J. Wang, J. Chen, J. Liu, and Z. Gong, "Trust-enhanced collaborative filtering for personalized point of interests recommendation," *IEEE Trans. Ind. Informat.*, early access, Dec. 29, 2019, doi: 10.1109/TII.2019.2958696.

[46] Z. Ning, P. Dong, J. J. P. C. Rodrigues, F. Xia, and X. Wang, "Deep reinforcement learning for vehicular edge computing: An intelligent offloading system," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 6, pp. 1–24, 2019.

[47] Z. Ning, Y. Li, P. Dong, X. Wang, M. S. Obaidat, X. Hu, L. Guo, Y. Guo, J. Huang, and B. Hu, "When deep reinforcement learning meets 5G-enabled vehicular networks: A distributed offloading framework for traffic big data," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1352–1361, Feb. 2020.

[48] Z. Ning, R. Y. K. Kwok, K. Zhang, X. Wang, M. S. Obaidat, L. Guo, X. Hu, B. Hu, Y. Guo, and B. Sadoun, "Joint computing and caching in 5G-envisioned Internet of Vehicles: A deep reinforcement learning-based traffic control system," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 5, 2020, doi: 10.1109/TITS.2020.2970276.

[49] Z. Ning, Y. Feng, M. Collotta, X. Kong, X. Wang, L. Guo, X. Hu, and B. Hu, "Deep learning in edge of vehicles: Exploring trirelationship for data transmission," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5737–5746, Oct. 2019.

[50] K. Zhang, Y. Zhu, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Deep learning empowered task offloading for mobile edge computing in urban informatics," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7635–7647, Oct. 2019.

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

IEEE *Access*

[51] W. Wei, X. Xia, M. Wozniak, X. Fan, R. Damaševi ius, and Y. Li, "Multi-sink distributed power control algorithm for cyber-physical-systems in coal mine tunnels," *Comput. Netw.*, vol. 161, pp. 210–219, Oct. 2019.

[52] P. Lin, Q. Song, J. Song, A. Jamalipour, and F. R. Yu, "Cooperative caching and transmission in CoMP-integrated cellular networks using reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5508–5520, May 2020.

[53] M. G. R. Alam, S. F. Abedin, S. I. Moon, A. Talukder, and C. S. Hong, "Healthcare IoT-based affective state mining using a deep convolutional neural network," *IEEE Access*, vol. 7, pp. 75189–75202, 2019.

[54] O. Iliashenko, Z. Bikkulova, and A. Dubgorn, "Opportunities and challenges of artificial intelligence in healthcare," in *Proc. E3S Web Conf.*, vol. 110. Les Ulis, France: EDP Sciences, 2019, p. 02028.

[55] S. Chenthara, K. Ahmed, H. Wang, and F. Whittaker, "Security and privacy-preserving challenges of e-Health solutions in cloud computing," *IEEE Access*, vol. 7, pp. 74361–74382, 2019.

[56] P. K. Maganti and P. M. Chouragade, "Secure application for sharing health records using identity and attribute based cryptosystems in cloud environment," in *Proc. 3rd Int. Conf. Trends Electron. Informat. (ICOEI)*, Apr. 2019, pp. 220–223.

[57] X. Liu, W. Ma, and H. Cao, "Mbpa: A MediBchain-based privacy-preserving mutual authentication in TMIS for mobile medical cloud architecture," *IEEE Access*, vol. 7, pp. 149282–149298, 2019.

[58] J. Liu, H. Tang, R. Sun, X. Du, and M. Guizani, "Lightweight and privacy-preserving medical services access for healthcare cloud," *IEEE Access*, vol. 7, pp. 106951–106961, 2019.

[59] K. S. Gill, S. Saxena, and A. Sharma, "Taxonomy of security attacks on cloud environment: A case study on telemedicine," in *Proc. Amity Int. Conf. Artif. Intell. (AICAI)*, Feb. 2019, pp. 454–460.

[60] G. Marquez, H. Astudillo, and C. Taramasco, "Exploring security issues in telehealth systems," in *Proc. IEEE/ACM 1st Int. Workshop Softw. Eng. Healthcare (SEH)*, May 2019, pp. 65–72.

[61] R. Guo, H. Shi, D. Zheng, C. Jing, C. Zhuang, and Z. Wang, "Flexible and efficient blockchain-based ABE scheme with multi-authority for medical on demand in telemedicine system," *IEEE Access*, vol. 7, pp. 88012–88025, 2019.

[62] K. Zhang, Y. Zhu, S. Maharjan, and Y. Zhang, "Edge intelligence and blockchain empowered 5G beyond for the industrial Internet of Things," *IEEE Netw.*, vol. 33, no. 5, pp. 12–19, Sep. 2019.

[63] F. Wang, H. Zhu, X. Liu, R. Lu, J. Hua, H. Li, and H. Li, "Privacy-preserving collaborative model learning scheme for E-healthcare," *IEEE Access*, vol. 7, pp. 166054–166065, 2019.

[64] Y. Fan, Y. Liao, F. Li, S. Zhou, and G. Zhang, "Identity-based auditing for shared cloud data with efficient and secure sensitive information hiding," *IEEE Access*, vol. 7, pp. 114246–114260, 2019.

[65] S. Wang, D. Zhang, and Y. Zhang, "Blockchain-based personal health records sharing scheme with data integrity verifiable," *IEEE Access*, vol. 7, pp. 102887–102901, 2019.

[66] J. Xu, K. Xue, S. Li, H. Tian, J. Hong, P. Hong, and N. Yu, "Healthchain: A blockchain-based privacy preserving scheme for large-scale health data," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8770–8781, Oct. 2019.

[67] H. Zhang, J. Yu, C. Tian, P. Zhao, G. Xu, and J. Lin, "Cloud storage for electronic health records based on secret sharing with verifiable reconstruction outsourcing," *IEEE Access*, vol. 6, pp. 40713–40722, 2018.

[68] X. Liu and W. Ma, "ETAP: Energy-efficient and traceable authentication protocol in mobile medical cloud architecture," *IEEE Access*, vol. 6, pp. 33513–33528, 2018.

[69] P. Lin, Q. Song, and A. Jamalipour, "Multidimensional cooperative caching in CoMP-integrated ultra-dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1977–1989, Mar. 2020.

[70] Z. Ning, X. Wang, and J. Huang, "Mobile edge computing-enabled 5G vehicular networks: Toward the integration of communication and computing," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 54–61, Mar. 2019.

[71] W. Shi, H. Sun, J. Cao, Q. Zhang, and W. Liu, "Edge computing-an emerging computing model for the Internet of everything era," *J. Comput. Res. Develop.*, vol. 54, no. 5, pp. 907–924, 2017.

[72] X. Liu, T. Qiu, X. Zhou, T. Wang, L. Yang, and V. Chang, "Latency-aware path planning for disconnected sensor networks with mobile sinks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 350–361, Jan. 2020.

[73] J. Ding and D. Fan, "Edge computing for terminal query based on IoT," in *Proc. IEEE Int. Conf. Smart Internet Things (SmartIoT)*, Aug. 2019, pp. 70–76.

[74] Z. Ning, J. Huang, X. Wang, J. J. P. C. Rodrigues, and L. Guo, "Mobile edge computing-enabled Internet of Vehicles: Toward energy-efficient scheduling," *IEEE Netw.*, vol. 33, no. 5, pp. 198–205, Sep. 2019.

[75] H. Chen, T. Zhao, C. Li, and Y. Guo, "Green Internet of Vehicles: Architecture, enabling technologies, and applications," *IEEE Access*, vol. 7, pp. 179185–179198, 2019.

[76] X. Yang and R. Bi, "Budget-aware equilibrium offloading for mobile edge computing," in *Proc. IEEE Int. Conf. Smart Internet Things (SmartIoT)*, Aug. 2019, pp. 383–387.

[77] Z. Ning, L. Liu, F. Xia, B. Jedari, I. Lee, and W. Zhang, "CAIS: A copy adjustable incentive scheme in community-based socially aware networking," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3406–3419, Apr. 2017.

[78] X. Wang, Z. Ning, and L. Wang, "Offloading in Internet of Vehicles: A fog-enabled real-time traffic management system," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4568–4578, Oct. 2018.

[79] Z. Ning, J. Huang, and X. Wang, "Vehicular fog computing: Enabling real-time traffic management for smart cities," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 87–93, Feb. 2019.

[80] K. Dolui and S. K. Datta, "Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing," in *Proc. Global Internet Things Summit (GIoTS)*, Jun. 2017, pp. 1–6.

[81] Z. Ning, X. Wang, J. J. P. C. Rodrigues, and F. Xia, "Joint computation offloading, power allocation, and channel assignment for 5G-enabled traffic management systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3058–3067, May 2019.

[82] W. Ni, H. Tian, X. Lyu, and S. Fan, "Service-dependent task offloading for multiuser mobile edge computing system," *Electron. Lett.*, vol. 55, no. 15, pp. 839–841, Jul. 2019.

[83] H. Guo, J. Liu, and J. Zhang, "Efficient computation offloading for multi-access edge computing in 5G HetNets," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[84] J. Li, H. Zhang, H. Ji, and X. Li, "Joint computation offloading and service caching for MEC in multi-access networks," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–6.

[85] M. Hua, H. Tian, W. Ni, and S. Fan, "Energy efficient task offloading in NOMA-based mobile edge computing system," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–7.

[86] Y. Yang, Y. Wang, R. Wang, and S. Chu, "A resource allocation method based on the core server in the collaborative space for mobile edge computing," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2018, pp. 568–572.

[87] H. Hui, C. Zhou, X. An, and F. Lin, "A new resource allocation mechanism for security of mobile edge computing system," *IEEE Access*, vol. 7, pp. 116886–116899, 2019.

[88] R. K. Pathinarupothi, P. Durga, and E. S. Rangan, "IoT-based smart edge for global health: Remote monitoring with severity detection and alerts transmission," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2449–2462, Apr. 2019.

[89] J. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, "Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks," in *Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2019, pp. 601–604.

[90] A. Emam, A. A. Abdellatif, A. Mohamed, and K. A. Harras, "Edge-Health: An energy-efficient edge-based remote mHealth monitoring system," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–7.

[91] P. H. Vilela, J. J. P. C. Rodrigues, P. Solic, K. Saleem, and V. Furtado, "Performance evaluation of a fog-assisted IoT solution for e-Health applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 379–386, Aug. 2019.

[92] A. Alabdulatif, I. Khalil, X. Yi, and M. Guizani, "Secure edge of things for smart healthcare surveillance framework," *IEEE Access*, vol. 7, pp. 31010–31021, 2019.

IEEE *Access*

L. Sun *et al.*: Edge-Cloud Computing and Artificial Intelligence in Internet of Medical Things: Architecture, Technology and Application

LANFANG SUN received the B.Sc. degree in network engineering from the Tianjin University of Technology, Tianjin, China, in 2019. She is currently pursuing the M.Sc. degree with the School of Software, Dalian University of Technology, Dalian, China. Her research interests include edge computing, medical Internet of Things, and big data.

HUIXIA REN received the Ph.D. degree from the University of Macau, Macao, in 2018. She is currently working with the Neurology department, The Second Clinical Medical College (Shenzhen People's Hospital), Jinan University for her Postdoctoral Program. She has authored or coauthored more than 20 scientific articles in her research interests which include unsaturated fatty acids, dementia especially Alzheimer's disease, and functional brain networks.

XIN JIANG received the Ph.D. degree from the Tongji Medical College, Huazhong University of Science and Technology, in 2008. From 2009 to 2010, she was a Visiting Scholar with the Prince of Wales Hospital, Chinese University of Hong Kong. She studied at the Global Clinical Scholars Research Training Program, from 2015 to 2016. She is currently the Chief Physician, an Associate Professor, and the Deputy Director of geriatrics with Jinan University.

YI GUO received the medical and M.Med. degrees from the Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, in 1985 and 1991, respectively, and the Ph.D. degree from the University of Greifswald, Greifswald, Germany, in 1997. He is currently the Chief of Neurology with The Second Clinical Medical College, Jinan University, a member of the cerebrovascular disease group of Chinese Medical Association neurology Branch, the Chairman of the Shenzhen Medical Association of neurology, and the Shenzhen Medical Association of Psychosomatic Medicine. His major research areas are cerebrovascular diseases, dementia, movement disorder diseases, sleep disorder, depression, and anxiety.

• • •