Cyber Security and Applications 1 (2023) 100010

Contents lists available at ScienceDirect

Cyber Security and Applications

journal homepage: http://www.keaipublishing.com/en/journals/cyber-security-and-applications/

Edge intelligent collaborative privacy protection solution for smart medical

Jinshan Lai [a,*], Xiaotong Song [a], Ruijin Wang [a], Xiong Li [b]

[a] School of information and software engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[b] School of Computer Science and Engineering, University and Technology of china, Chengdu 611731, China

ARTICLE INFO

Keywords:
Edge intelligent collaboration
Privacy protection
Smart medical care
Differential privacy

ABSTRACT

In the era of big data, competent medical care has entered people's lives. However, the existing intelligent diagnosis models have low accuracy and poor universality. At the same time, there is a risk of privacy leakage in the process of health monitoring and auxiliary diagnosis. This paper combines edge computing and federated learning ensure model accuracy and protect patient privacy by proposing an Edge intelligent collaborative privacy protection solution for smart medical (EICPP). First, we offer a lightweight edge intellectual collaborative federated learning framework named KubeFL to support health monitoring and auxiliary diagnosis; secondly, we design a federated learning training model based on device-edge-cloud layering, with complete accuracy of up to 95.8%; Finally, a differential privacy algorithm for edge-cloud model transmission is proposed, which can exchange a lower accuracy loss for solid privacy protection.

1. Introduction

With the increasing number of older adults, the health monitoring and disease diagnosis and treatment of the elderly living alone have become a difficult complex problem. For example, the elderly cannot get timely help for a fall or a sudden heart attack at home, and the elderly cannot move to the hospital in time to check whether they have diabetes or other diseases. With the vigorous development of new-generation information technologies such as 5G, big data, cloud computing, Internet of Things, edge computing, and federated learning, the combination of digital, networked, and intelligent facilities and solutions with medical scenarios has enabled innovative medical applications in AI-assisted diagnosis, health management, telemedicine, and other fields. Tencent launched the "Dr. Clove" app [1], which includes three functions: online consultation, online drug purchase, and health science popularization, and is committed to providing reliable medical and health information and services. For some acute diseases or public health emergencies, such as stroke, new coronary pneumonia (Corona Virus Disease 2019, COVID-19), etc., patients can obtain timely and correct treatment advice through telemedicine, to buy time for the treatment of the disease [2,3].

It is a great responsibility to solve the problems of elderly health care and disease diagnosis and treatment. Smart medical care is the key to solving this problem. The use of wearable devices and mobile phones, combined with advanced technologies such as artificial intelligence, can realize functions such as health monitoring and online diagnosis. However, in practice, in the process of real-time monitoring and online diagnosis of the elderly combined with competent medical care, the following three problems are often encountered:

1) The elderly cannot quickly and accurately notify their family members in case of an accident;
2) Online intelligent diagnosis has low accuracy and poor universality;
3) There is a risk of privacy leakage in the process of health monitoring and auxiliary diagnosis;

Specifically, smart medical care takes medical cloud data as the core, uses the Internet of Things and data transmission and exchange as technologies, and combines electronic medical records and electronic health files to build a medical system for medical and health services [4]. However, medical data is extremely privacy-sensitive. Suppose the user's original sign data is directly submitted to the cloud data center. In that case, there is a risk of privacy leakage in communication, processing, and storage, such as man-in-the-middle attacks and unintentional viewing by cloud data center managers or malicious transactions. To this end, federated learning is widely used as a key technology for privacy protection [5]. Federated learning realizes the model training process in which data is not local. Still, the computing power of the mobile medical devices is often limited. Complete neural network training cannot be performed, so it is necessary to use the architecture of edge computing [6,7] to offload most of the computing tasks to the edge server [8–10]. However, it is also unreliable to transmit all raw data of mobile medical devices to edge servers because edge servers face more security and privacy threats than cloud data centers. In addition, the models trans-

* Corresponding author.
E-mail addresses: 958101695@qq.com (J. Lai), 1098304107@qq.com (X. Song), ruijinwang@uestc.edu.cn (R. Wang), lixiongzhq@163.com (X. Li).

https://doi.org/10.1016/j.csa.2022.100010
Received 1 July 2022; Received in revised form 4 September 2022; Accepted 4 October 2022
Available online 13 October 2022

mitted in the federated learning process also face the threat of inference attacks, and these problems must be solved well [11].

Although federated learning does not exchange data directly and has a higher privacy guarantee than traditional centralized machine learning training, there are still two problems:

1) Federated learning itself does not provide comprehensive and adequate privacy protection. It still faces the threat of information leakage. For example, attack methods can already infer user data based on the intermediate parameters $\bigtriangledown W = \partial \text{Loss}/\partial \text{W}$ of model training.
2) Terminal devices are often unable to efficiently and smoothly complete local model training tasks, especially mobile medical devices in medical scenarios, with limited computing and storage resources, making it challenging to complete the training of massive neural networks.

In response to the above problems, we propose an edge intelligent collaborative privacy protection scheme (EICPP) for competent medical care and offer a lightweight edge intellectual collaborative federated learning framework KubeFL, which coordinates various devices and technologies from the overall architecture to realize edge brilliant collaboration Model training in computing mode; use the federated learning training model based on the device-edge-cloud layering, follow the principle that the original data is not local, and efficiently, accurately and safely combine the data collected by each device to help determine whether the elderly fall Whether you may suffer from various chronic diseases, be prepared for prevention. The differential privacy algorithm for edge-cloud model transmission is used to protect user privacy so that attackers with ulterior motives cannot obtain user information and ensure the information security of the entire process.

The KubeFL framework can achieve efficient management of different entities, efficient and convenient health monitoring, and efficient and timely auxiliary diagnosis; in terms of model accuracy, accurate sign recognition can be achieved, and after testing, the total recognition success rate is as high as 95%; in addition, in terms of privacy protection, the federated learning training model based on end-edge-cloud layering is used, and the original data is not local, which reduces the risk of user privacy leakage at the root. Finally, differential privacy is used in uploading the model to the cloud data center and sending it to the edge serverso that the attacker cannot obtain the detailed data of the training through inference attacks. The privacy and security of the model are guaranteed.

2. Related work

2.1. Federated learning

Traditional machine learning and deep learning require centralized training by service providers after collecting user data. However, user data is closely related to individual users and may directly or indirectly contain sensitive information. If the server provider leaks such sensitive information, it will now threaten the safety of the person, reputation and property of users; another problem is that the data collected by a single service provider is limited, there are regional differences, and the generalized application of the model cannot be guaranteed, that is, there is the problem of data silos. In the context of the increasingly prominent contradiction between the phenomenon of data silos and data security requirements, federated learning emerges as the times require. The learning mechanism for sharing training parameters is only exchanged in the middle stage for raw data. Ideally, the shared parameters obtained by federated learning are similar or better than those obtained by training the data on the central server.

A typical architecture of federated learning for healthcare is shown in Fig. 1.

2.2. Differential privacy

Differential privacy is an approach in cryptography that aims to maximize the accuracy of data queries when queried from statistical databases while minimizing the chance of identifying their records. It has the following characteristics:

(1) Imagine a powerful attacker. If there are n people in the source data, the attacker knows the information of $n-1$ people and tries to get the last person's information. In this case, the attacker still cannot get the previous person's information. Since such attackers hardly exist, differential privacy has a good protection effect.
(2) Based on rigorous data theory. Differential privacy technology can significantly facilitate mathematical tools, quantitative analysis, and proof, and has good scalability.
(3) Extremely high privacy protection performance. Compared with not using any privacy-preserving algorithm, the differential privacy algorithm uses less accuracy drop in exchange for highly high privacy-preserving performance.

The specific definitions are as follows.

If the difference between the two data sets D, $D^{'}$ is only reflected in a single sample, that is, only one record is different, then D, $D^{'}$ is called adjacent data sets. Given an algorithm M: $D \rightarrow R$, where the domain of M is D and the range is R, for any two adjacent datasets d, $d \in D$ and the output subset $s \in R$, there are inequalities.

$$Pr[M(d) \in S] \leq e^{\epsilon} Pr[M(d^{'}) \in S] + \delta \tag{1}$$

If established, the algorithm M satisfies (ϵ, δ)-differential privacy. ϵ is the privacy budget; the smaller the ϵ, the higher the privacy protection level, and the more noise is added. The lower the data availability, δ is the relaxation factor, indicating the probability that the above inequality is not satisfied, then $\delta = 0$ called ϵ-differential privacy.

Differential privacy methods are usually implemented by adding fuzzy noise, and their essence is to protect sensitive information by adding quantitative randomness to the data. Commonly used noise-adding mechanisms are the Gaussian mechanism and Laplace mechanism. In the Gaussian mechanism, we have the following formula.

$$M(d) \triangleq f(d) + N(0, S_f^2 * \sigma^2) \tag{2}$$

where $M(d)$ represents the result after adding noise, $N(0, S_f^2 * \sigma^2)$ is a normalized distribution with a mean of 0, a standard deviation of $S_f * \sigma$ and S_f is the sensitivity of the function f.

$$S_f = \max_{(d,d')} ||f(D) - f(D^{'})||_2 \tag{3}$$

By adding noise to the function f, a new function M is constructed, and it can be proved that M also satisfies (ϵ, δ)-differential privacy.

3. Scheme

In this section, we describe the system model and our scheme.

3.1. System architecture

In our model, there are three entities.

1) Mobile medical devices: They can collect medical data of users and patients and then upload it to the hospital. Due to limited computing and storage resources, only simple data processing can be performed to train the model. Low efficiency and inaccurate training results;
2) Hospital servers: servers distributed in various places can perform computing tasks according to the requirements of hospitals and users but that cannot guarantee the security of internal data, there is a risk of data leakage, and due to the geographical location of hospitals due to differences, different user groups and the characteristics of the hospital itself, the information available to each hospital is limited and biased.

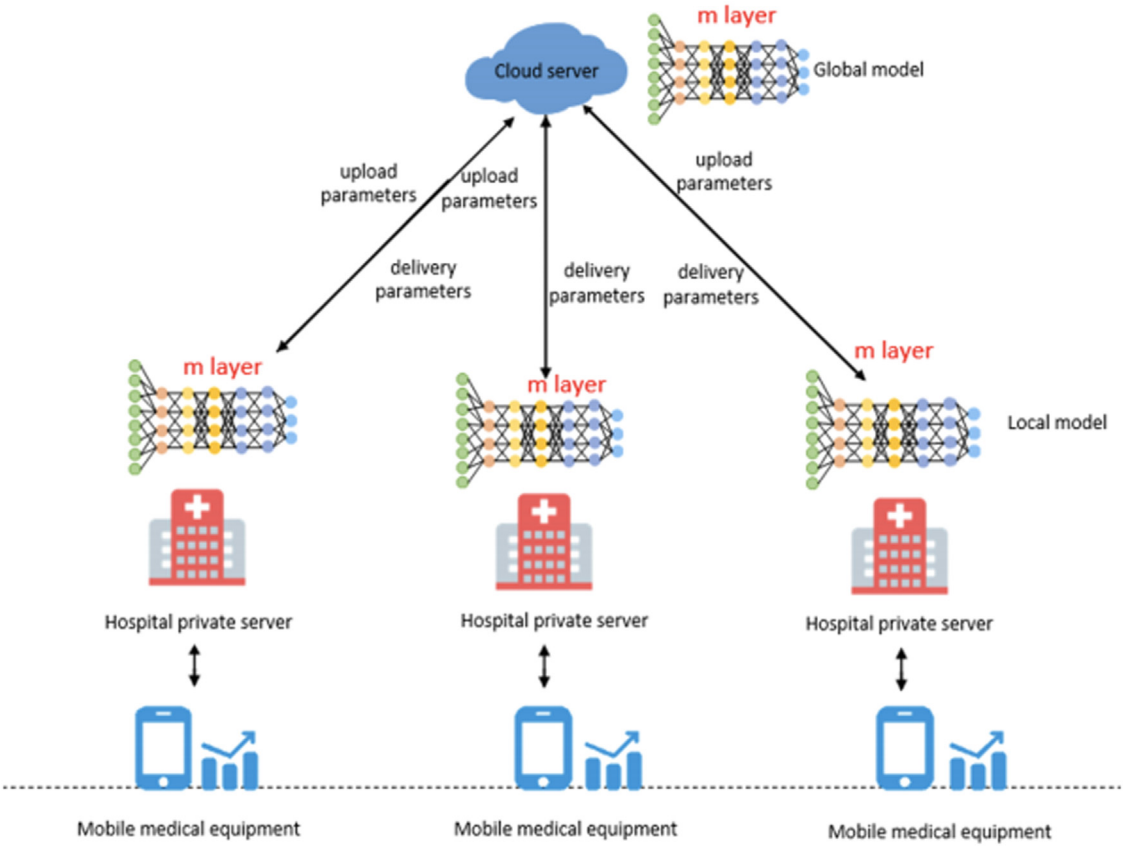

Fig. 1. Federated learning framework.

3) Cloud data center: The cloud data center is a centralized platform that can collect data from various hospitals. It has strong security protection and rich computing and storage resources. However, it is far from users and inconvenient to communicate with them.

The privacy leakage problem in the federated learning training process can be solved by differential privacy technology, but it also increases the computational burden of the client device. If the data is uploaded to the private server of the hospital, there will be a data leakage problem. If a medical device collects data and completes the model training task yourself and then uploads the model parameters to the private server of the hospital, although the problem of data leakage can be prevented due to the limited computing and storage resources of mobile medical devices, there will be problems of training efficiency and accuracy. Considering the above issues, our scheme offloads the first part of the training model task to the medical device, and the second part is placed on the private server of the hospital. The cloud center is responsible for aggregating the model and storing the model so that the data of the end device is not uploaded to a third party, and the model training task is also processed. Successful completion can protect the original data of users and patients from leaking. The overall framework is shown in Fig. 2.

3.2. Training process

The specific workflow of a complete medical model training is shown in Fig. 3, which mainly includes the following processes:

1) The hospital collects data from the medical records or wearable devices of users and patients;
2) The hospital checks whether the collected data meets the model training requirements ;
3) If the requirements are met for model training, if not, continue to collect data until the requirements are met;
4) After meeting the requirements, upload the data to the hospital server for model training;
5) Upload and store the trained model to the cloud center server; the model can be obtained by accessing the cloud center server when the model is needed in the future.

The training algorithm flow is as follows.

In Algorithm 1, we introduce the overall flow of training and the interaction between the three entities. As can be seen from Algorithm 2, the mobile terminal device trains part of the model and sends the training results to the hospital server, and at the same time, receives the gradient to update the model parameters. Then the hospital server trains the complete model, uploads the gradient to the cloud server, receives

Algorithm 1 Training process.

1: The central server distributes the initial model to many hospital private servers
2: The private server of the hospital splits the model and distributes the first half of the model to mobile medical devices
3: Mobile medical device collects user and patient data, which are divided into different groups $D_1, D_2 \ldots D_k$ $(D_i = (X_i, Y_i))$
4: These data are input into the designed deep neural network model
5: The mobile medical device sends the output results and labels of the front part of the model to the private server of the hospital
6: The private server of the hospital trains the remaining model and adds the differential disturbance to the result and sends it to the central cloud server
7: The central cloud server aggregates and updates the model and distributes it to the hospital server for the next round of training

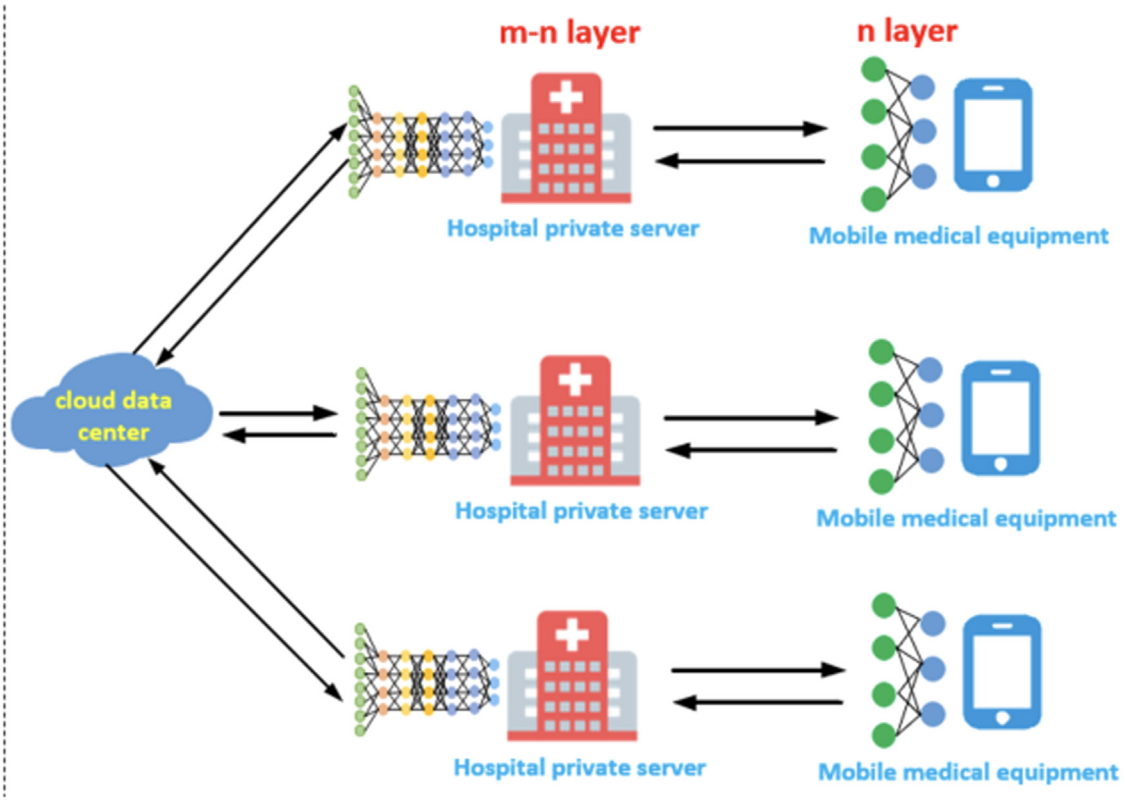

Fig. 2. Framework of federated learning based on device-edge-cloud hierarchy.

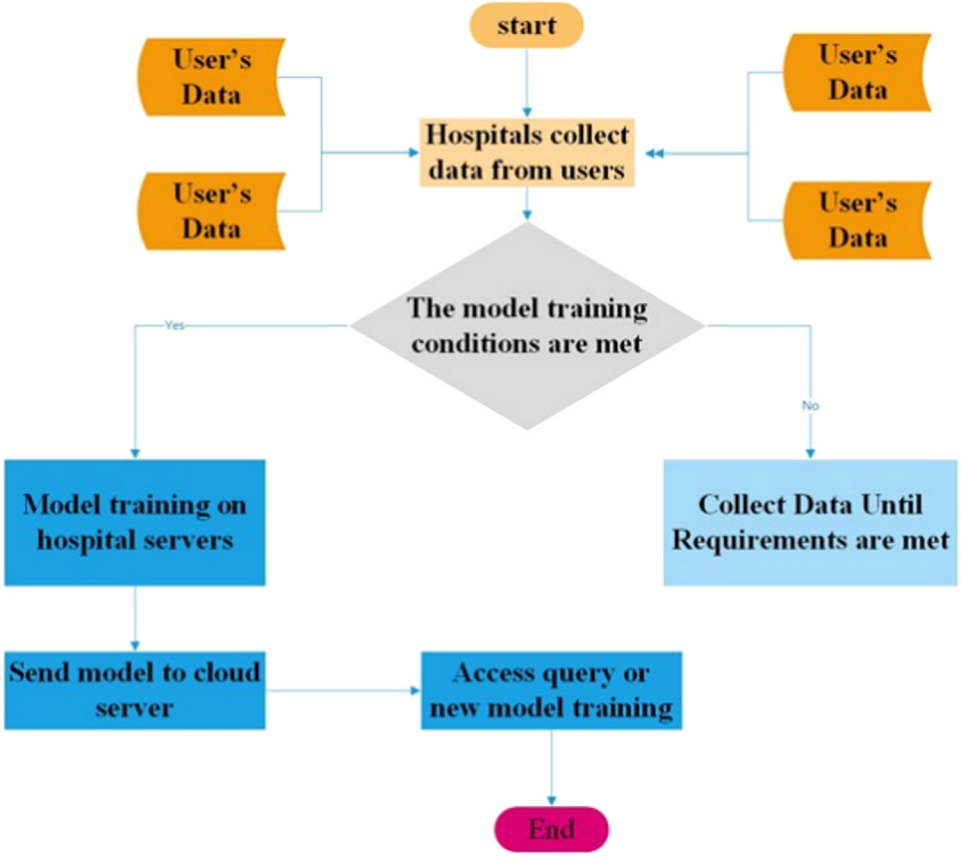

Fig. 3. Medical model training workflow.

Algorithm 2 Local training.

Initialize client model parameter W_{note}
for each medical device **do**
  Divide the data into k groups $D_1, D_2 \ldots D_k$
  for i in range t **do**
    Input W_{note} and D_i into the first n layers of neural network to get output O_n
    Add the noise to O_n to get O'_k
    Upload the labels Y_i of the O'_k and D_i data groups to the hospital's server
    Receive $\nabla Loss(O'_k) = \partial Loss / \partial(O'_k)$ from the hospital server
    Update parameter W_{note}
    $W_{note-new} = W_{note-old} - \eta Loss(O'_k) * \nabla O'_k(W_{note-old})$

the gradient aggregated by the cloud server to update the parameters, and sends the gradient to the mobile terminal device. In Algorithm 4, the cloud aggregates the gradients uploaded by the hospital server and adds differential noise, and then sends it back to the hospital server for the next round of training.

4. Experimental analysis

In this section, we use python to conduct simulation experiments, analyze the simulation experiments' accuracy, efficiency, and privacy, and compare and analyze with other federated learning schemes.

Algorithm 3 Parameter update of edge server.

for each dege server **do**
  Download the parameter W_{cloud} of the cloud center server global model
  Assign the parameter W_{cloud} of the global model of the cloud center server to the local model parameter $W_{hospital}$
  Receive the label Y_i of the O'_k and D_i data groups from the end device
  Update the parameters
  $W_{hospitalnew} = W_{hospitalold} - \eta \nabla Loss(W_{hospitalold})$
  Pass $\nabla Loss(O'_k)$ back to the end device until the required accuracy

Algorithm 4 Cloud aggregation algorithm.

Initial recognition of the global model parameter W_{cloud}
for i in range p **do**
  Randomly select hospitals from q hospitals into the set Z_p
  Assign $\triangle$W to the empty set
  for j in range Z_p **do**
    Pass W_{cloud} to hospital j
    The hospital obtains the updated parameter $W^j_{hospitalnew}$ through cooperative computing with the terminal device
    The cloud center server will receive $W^j_{hospitalnew}$ from the hospital's server
    After receiving $W^j_{hospitalnew}$, the cloud center server will compare the difference between $W^j_{hospitalnew}$ and W_{cloud}
    Update parameter $W^j_{diff} = W^j_{hospitalnew} - W_{cloud}$
    Update $\triangle W = \triangle W \cup W^j_{hospitalnew}$
    Add the noise to $\triangle$W to get $\triangle W'$
$W_{cloudnew} = W_{cloudold} + 1/(|Z_q|) \triangle W'$
Send $W_{cloudnew}$ to hostpital server

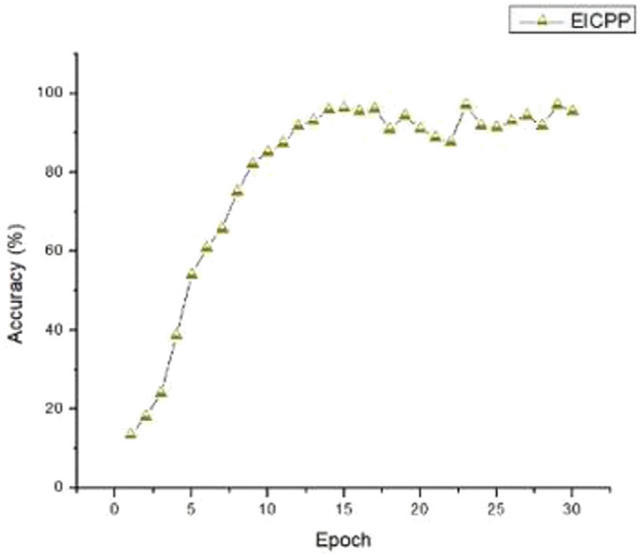

Fig. 4. Accuracy of EICPP on breast cancer dataset.

4.1. Experimental configuration

To verify the effect of EICPP in terms of accuracy and privacy protection, we use PyTorch to conduct simulation tests on the fully connected network using the breast cancer dataset [12]. To facilitate experiments, we assign the first layer of the model to medical devices for training and get the accuracy, recall, precision, and F1 values; at the same time, we compare with FedAvg and FedMA to get the comparison value of accuracy of different methods; adjust different privacy budgets at the same time, and test the accuracy of models under different separate privacy budgets variety.

Specifically, we use four Raspberry Pis with the ubuntu operating system as mobile medical devices, two PCs with the ubuntu operating system as hospital servers, and Trend Cloud as cloud servers. Among them, the memory of the Raspberry Pi is 2 GB, and the PC's memory is 4 GB.

4.2. Precision analysis

To test the precision of EICPP, we conduct experiments from several dimensions of accuracy (Accuracy), recall (Rec), precision (Pre), and harmonic mean (F1) and obtain the experimental results as shown below. As shown in Fig. 4, with the increase in training discussion, the model's accuracy increases rapidly, reaching more than 90% around the 12th round. After 30 rounds of training, the final accuracy rate can get 95.47%; it can be seen that the recall rate show the same upward trend and can finally go 0.955 from Fig. 5; similarly, as can be seen from Fig. 6, the precsion of the model can reach 0.959; from Fig. 7 can be obtained that the harmonic mean of the model can 0.957; from the above four dimensions, it can be seen that EICPP can achieve quite a high accuracy on the breast cancer data set, which is of great help for clinical medical diagnosis and treatment.

In addition, we also compare the accuracy with FedAvg and FedMA, using the breast cancer dataset. The obtained accuracy comparison chart is shown in Fig. 8. It can be seen from the figure that after 30 rounds, the accuracy of the model that under breast cancer dataset tends to convergethe. Specifically, accuracy rate growth rate of EICPP is higher than that of FedAvg and FedMA. EICPP is close to convergence around the 14th round, FedMA is close to confluence around the 20th round, and FedAvg is only relative to the 24th round reaching a junction. At the same time, from comparing the accuracy rate at the final intersection, it can be seen that EICPP can get 95.47%, FedMA can go 82.56%, and

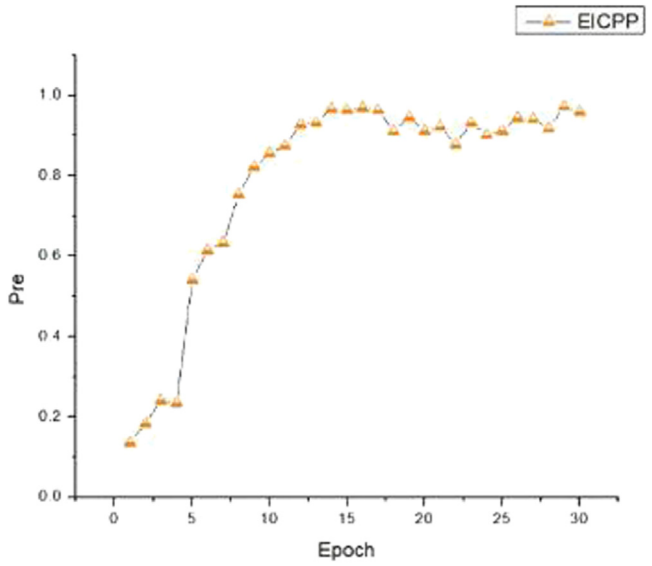

Fig. 5. Pre of EICPP on breast cancer dataset.

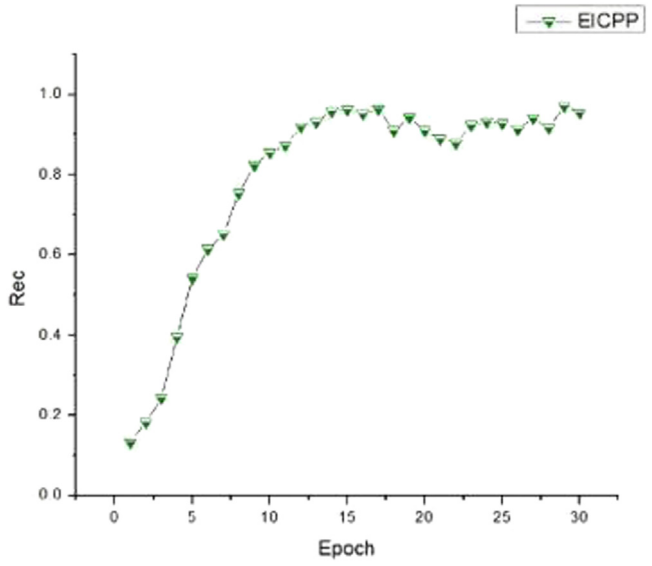

Fig. 6. Rec of EICPP on breast cancer dataset.

Table 1
Comparison of EICPP and PPFLEC .

Schemes	PPFLEC	EICPP
Accuracy of Mnist datasets	91.52%	95.36%
Accuracy of CIFAER10 datasets	65.5%	67.2%

FedAvg can only reach 70.35%. It can be seen that the accuracy of EICPP is much higher than the existing federated learning scheme. In medical diagnosis scenarios that require high precision.

In addition, we compare the accuracy of EICPP and PPFLEC [11] on the Mnist dataset [13] and Cifar10 dataset [14], the latter is a privacy protection scheme under "cloud-edge-end," and obtained the accuracy comparison chart as shown in Table 1.

Under the two datasets, the accuracy rates of EICPP are 95.36% and 67.2%, respectively, while PPFLEC is only 91.52% and 65.5%. We can

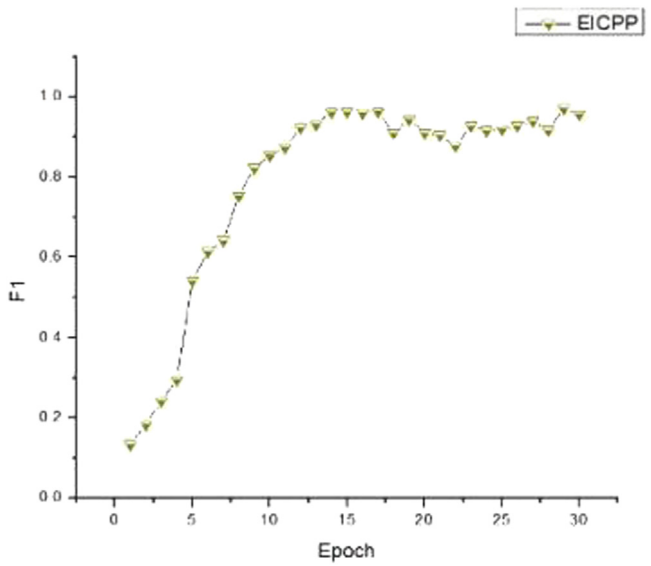

Fig. 7. F1 of EICPP on breast cancer dataset.

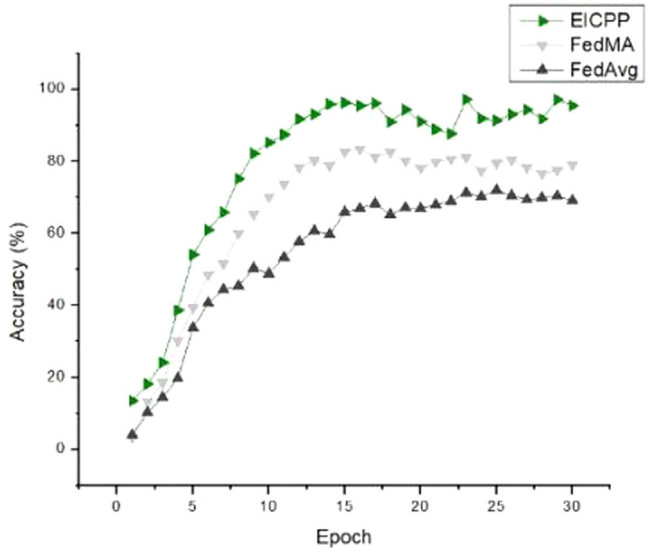

Fig. 8. Comparison of the accuracy rates of the three schemes under breast cancer data.

Table 2
Convergence time comparison table of three schemes .

Schemes	EICPP	FedMA	FedAvg
convergence time(s)	1050	1300	1600

conclude that EICPP has a higher accuracy rate and is more suitable for high precision and privacy protection. Edge computing scenarios.

4.3. Efficiency analysis

To test the convergence efficiency of EICPP, we train the three schemes on the same device and record the time required for the model to converge, as shown in Table 2.

As can be seen from Table 2, FedAvg has the longest convergence time, while FedMA has made improvements on FedAvg, and the time is shortened by about 300 s, while the EICPP model takes the least convergence time, only 1050 s, which is 550 s less than FedAvg. FedMA is

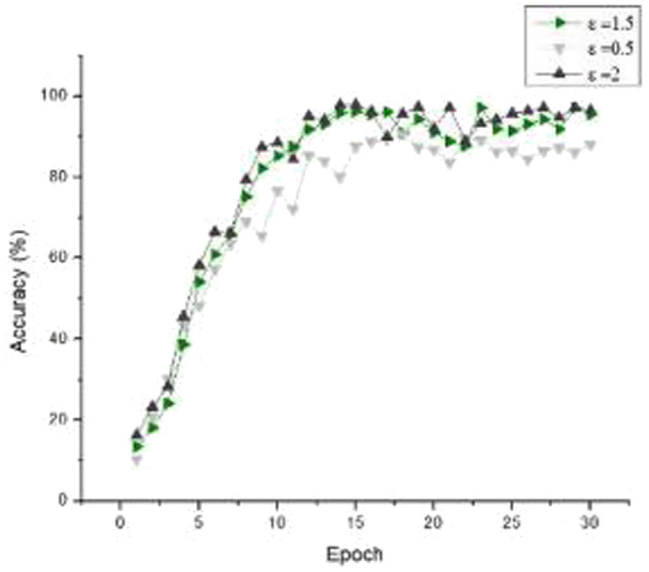

Fig. 9. Comparison chart of model accuracy under different privacy budgets.

250 s less, which shows that it has high efficiency in model convergence and can be applied to real-time medical scenarios.

4.4. Privacy analysis

To test the impact of different privacy budgets on the accuracy of EICPP, we set different privacy budgets to test the accuracy of EICPP under different privacy budgets, as shown in Fig. 9.

As seen from Fig. 9, when the privacy budget is 2, the model has the highest accuracy when it converges.When the privacy budget is 0.5, the accuracy is the lowest, which proves that the larger the privacy budget, the most negligible impact on the model accuracy. The higher the accuracy rate however, the privacy protection effect will weaken when the privacy budget is too large. Therefore, it is necessary to evaluate the relationship between the model accuracy rate and the privacy protection degree for different scenarios to set the privacy budget to achieve the desired effect.

To sum up, EICPP can protect patients' privacy in medical scenarios. At the same time, it can achieve a high model accuracy and realize intelligent diagnosis and treatment. Further, EICPP can not only be applied in medical scenarios; other scenarios that need to balance both privacy and accuracy can be used, such as intelligent model training in electricity and banking.

5. Conclusion and future work

This paper proposes an edge intelligent collaborative privacy protection scheme for competent medical care with high accuracy and training efficiency. At the same time, parameters can be adjusted according to different scenarios to achieve different degrees of privacy protection, which can well solve the current problems faced by medical diagnoses-issues such as low accuracy, low training efficiency, and data leakage. In the future, we will study the impact of different model partitions on model accuracy and improve our KubeFL framework to improve model accuracy.

Declaration of Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgment

This work is supported by Sichuan Science and Technology Program (2020YFG0475, 2018GZ0087, 2019YJ0543, 22DZX0046).

References

[1] G. Nittari, R. Khuman, S. Baldoni, G. Pallotta, G. Battineni, A. Sirignano, F. Amenta, G. Ricci, Telemedicine practice: review of the current ethical and legal challenges, Telemed. e-Health 26 (12) (2020) 1427–1437.
[2] A.E. Loeb, S.S. Rao, J.R. Ficke, C.D. Morris, L.H. Riley III, A.S. Levin, Departmental experience and lessons learned with accelerated introduction of telemedicine during the COVID-19 crisis, J. Am. Acad. Orthop. Surg. (2020).
[3] Y. Li, Y. Song, W. Zhao, X. Guo, X. Ju, D. Vogel, Exploring the role of online health community information in patients decisions to switch from online to offline medical services, Int. J. Med. Inform. 130 (2019) 103951.
[4] M. Leyva-Vázquez, M.A. Quiroz-Martínez, Y. Portilla-Castell, J.R. Hechavarría-Hernández, E. González-Caballero, A new model for the selection of information technology project in a neutrosophic environment, Neutrosophic Sets Syst. 32 (2020) 343.
[5] X. Xia, F. Chen, Q. He, J. Grundy, M. Abdelrazek, H. Jin, Online collaborative data caching in edge computing, IEEE Trans. Parallel Distrib. Syst. 32 (2) (2020) 281–294.
[6] M. Zwolenski, L. Weatherill, The digital universe: rich data and the increasing value of the internet of things, J. Telecommun. Digit. Econ. 2 (3) (2014). 47-1
[7] P. Bellavista, L. Foschini, D. Scotece, Converging mobile edge computing, fog computing, and IoT quality requirements, in: 2017 IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud), IEEE, 2017, pp. 313–320.
[8] S. Zhang, P. He, K. Suto, P. Yang, L. Zhao, X. Shen, Cooperative edge caching in user-centric clustered mobile networks, IEEE Trans. Mob. Comput. 17 (8) (2017) 1791–1805.
[9] Y. Zhao, W. Wang, Y. Li, C.C. Meixner, M. Tornatore, J. Zhang, Edge computing and networking: a survey on infrastructures and applications, IEEE Access 7 (2019) 101213–101230.
[10] J. Qadir, B. Sainz-De-Abajo, A. Khan, B. Garcia-Zapirain, I. De La Torre-Diez, H. Mahmood, Towards mobile edge computing: taxonomy, challenges, applications and future realms, IEEE Access 8 (2020) 189129–189162.
[11] R. Wang, J. Lai, Z. Zhang, X. Li, P. Vijayakumar, M. Karuppiah, Privacy-preserving federated learning for internet of medical things under edge computing, IEEE J. Biomed. Health Inform. (2022).
[12] Z. Mushtaq, A. Yaqub, S. Sani, A. Khalid, Effective *k*-nearest neighbor classifications for wisconsin breast cancer data sets, J. Chin. Inst. Eng. 43 (1) (2020) 80–92.
[13] L. Deng, The MNIST database of handwritten digit images for machine learning research [best of the web], IEEE Signal Process. Mag. 29 (6) (2012) 141–142.
[14] A. Krizhevsky, G. Hinton, et al., Learning multiple layers of features from tiny images (2009).