Automation and Machine Learning (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/autml.2025.060109
ISSN 2516-5003 Vol. 6 Num. 1

Intelligent recognition system based on federated learning

Lan Xiaoping, Niu Wenxue, Ge Lina, Wang Zhe

School of Artificial Intelligence, Guangxi University, Nanning, 530006, China

Keywords: Federated Learning; Intelligent Recognition System; Differential Privacy; Convolutional Neural Network; CNN; Data Island

Abstract: With the rapid development of deep learning, target detection algorithms have been widely used, but the traditional target detection methods need to collect a large number of labeled sensitive data, which is likely to violate user privacy and data confidentiality. As a privacy-preserving distributed machine learning method, federated learning enables end-to-end computer vision tasks, where image annotation and training tasks are moved to the edge, while only model parameters are sent to the aggregation server for aggregation. This paper proposes a kind of edge auxiliary iot intelligent recognition based on federal learning system, the system adopts the terminal layer, edge service layer, network layer and cloud center service layer four layer architecture, can analyze the distribution of detailed statistics, in the way of privacy protection, auxiliary iot devices for safe and intelligent object recognition.

1. Introduction

With the rapid development of artificial intelligence (Artificial Intelligence, AI) technology, machine learning and deep learning have become the core technology [1] in the field of computer vision. The main goal of machine learning is to extract useful features from image data through algorithms and models, and classify, identify or analyze images. Currently, it has been used in [2] fields such as face recognition, object detection and recognition, image classification, image segmentation, image generation, medical image analysis and intelligent transportation. Among them, deep learning [3] is a branch of machine learning, which is mainly based on the artificial neural network model, has a multi-level structure, and plays an important role in the field of image recognition. Image recognition refers to the identification and classification of objects, scenes, or features in an image from the input image data. Deep learning builds a deep neural network model and uses a large amount of annotated data for training, so as to realize efficient image recognition [4]. In deep learning, convolutional neural network (CNN) is one of the most commonly used architectures, especially suitable for image recognition tasks. CNN is able to effectively capture local features in the image and identify objects in the image through multi-level abstract expression. Through the back propagation algorithm, CNN is able to automatically learn the feature representation in the image and classify or identify them according to these features. The continuous development and optimization of deep learning has greatly improved image recognition in both accuracy and efficiency, bringing new applications and solutions of [5] to many fields. However, the most advanced deep learning intelligent models are becoming more and more large, containing billions or even trillions of learnable parameters. High-dimensional neural networks and

additional security solutions lead to deep learning models facing huge communication overhead, which has become the main bottleneck for the scale expansion of deep learning models. At the same time, most sensor devices, network storage devices and other devices process and store massive data all the time. The distributed storage of massive user privacy data makes it difficult for artificial intelligence models to use the data and experience of all parties, and it is difficult to guarantee the privacy security issues in the process of model sharing [6]. At present, the shortcomings restricting the development of deep learning technology in the field of image recognition involve many aspects.

Data security. Untrusted malicious adversaries may send incorrect model updates to the parameter aggregation server to compromise the model accuracy. The training process of deep learning model still involves the threat of data security attack, and may be subject to various forms of attack means, such as poisoning attack, back door attack, hiking attack, etc.

Privacy leakage: a malicious client or semi-honest learning party may launch a model inference attack to detect the model parameter information, so as to infer the local data set, resulting in local privacy data leakage.

Data island: At present, Internet of Things sensing devices collect rich data types with strong heterogeneity, and due to privacy security issues, the data sharing between different fields is not high, the degree of data integration is low, and the data is relatively scattered. The data between different fields forms a "data island", which is difficult to integrate and utilize, and is not conducive to the deep learning model with more perfect training function.

In order to effectively solve the problems of data "islands" and privacy security facing machine learning and deep learning, Google proposed the federated Learning (Federated learning, FL) [7] in 2016. As a paradigm for train statistical learning models in a distributed edge network, FL allows participants to train the model without sharing their local data, effectively alleviating data security risks and breaking the data "island" [8]. As a distributed machine learning method, FL implements the global model training without the source data. FL can better integrate decentralized data resources and conduct in-depth analysis, facilitate the formulation of real-time decision-making and situation perception, effectively make use of decentralized data resources, and contribute to the development of [9] in the field of image recognition. Federal learning has been widely used in health care [10], smart agriculture [11], smart city [12], smart industry [13] and other fields. In view of the challenges of current deep learning-based image processing technologies, this paper presents an edge-assisted IoT intelligent recognition system based on federated learning. The main contributions of this paper are as follows.

First, this paper designs a federated learning-based intelligent recognition system using the four-layer architecture of terminal layer, edge service layer, network layer, network layer and cloud center service layer, to assist secure and intelligent object recognition in a privacy-protected way. Secondly, the handwritten image MNIST dataset is used to verify the effectiveness and accuracy of the proposed federated learning-based intelligent recognition system. The results show that the proposed system can achieve high model accuracy by privacy protection.

2. FL overview with related techniques

2.1 FL definition

Federated learning aims to build a federated learning model with distributed datasets. Figure 1 shows an example of a federated learning architecture with coordinating ators. This scenario, the cloud parameter server as coordinator, can $\{D_i\}_{i=1}^{n}$ send the initial model to N participants, each participant using their respective training data set collaborative training machine learning model,

and update the model weight encryption parameters uploaded to the parameter server, parameter server will receive the model aggregation, aggregation method can be the federal average algorithm. Later, the parameter server sends the aggregated model updates back to the participant. This process is performed repeatedly until the model converges or the maximum number of iterations is reached. Federated learning architecture can be designed as peer-to-peer networks. Under this architecture, each participant is in a reciprocal position with no subordinate. Each participant can act as a model training party, provide model parameters for other participants, and participate in federal modeling and federal reasoning. It can also act as a business party to initiate the federal modeling and federal reasoning of business scenarios. This architecture eliminates the hidden danger of a single point of failure problem, further ensuring the system security, and is easy to scale, but may consume more computing resources in the encryption and decryption of message communication [14].

2.2 Classification of the federated learning

Suppose the data of the m th participant is represented by a matrix Dm, each row of the matrix Dm represents a data sample, and each column represents a specific data feature. For datasets with label information, the assumed feature space is X, data label space is Y, and sample ID space is I. According to the distribution of training data in the data feature space and sample ID space between different participants, federated learning is divided into horizontal federated learning, and longitudinal federated learning and federated transfer learning.

2.2.1 Transverse federal learning

Horizontal federated learning is applicable where the data characteristics of the participants overlap more. For example, as shown in Table 1, when the federal learning participants A and participant B are two different banks, they may have only A few overlapping customers, but different customers I1, I2,..., in's data may have very similar characteristics space because of similar business model, namely the two bank users overlap smaller, but the data characteristics overlap is larger, the two Banks can cooperate through horizontal federal learning to build a machine learning model. Therefore, lateral federated learning is also known as federated learning by sample.

Table 1. Horizontal federal learning

	ID	X1	X2	X3	Y
participant A	I1				
	I2				
participant B	I3				
	I4				

2.2.2 Longitudinal federal learning

Longitudinal federated learning is characterized by aligned data samples between participants, but varying in data characteristics. For example, as shown in Table 2, participant A and participant B provide different services, but when there are very large intersections in customer groups, they can collaborate on their own different feature spaces to get a better statistical learning model for each. Due to the large overlap between users, the data features are small. Therefore, longitudinal federated learning is also called federated learning by characteristics.

Table 2: Longitudinal federal learning

participant A	ID	X1	X2	Y	participant B	ID	X3	X4	Y
	I1								
	I2								

2.2.3 Federal Transfer Learning

Federated transfer learning allows the transfer of complementary knowledge across domains in data federation, with small modifications to existing model structures that help applications with small data (less overlapping samples and features) and weakly supervised (less labeled) to build effective and accurate machine learning models. Especially in medical, intelligent manufacturing, data prediction, graphics processing and other data heterogeneity is particularly strong application fields. Federated transfer learning is based on the data distributed in multiple parties. It is set for the data of different participants not only different sample space but also different feature space. It is applicable to the data samples and data characteristics of federated learning participants have little overlap. The integration of the transfer learning in the federal learning system may largely solve the limitations of the transfer learning application, which is very beneficial to the improvement and development of the federal learning system. Federated transfer learning is characterized by very little overlap between the data samples and the data characteristics of the participants. Federated transfer learning is modeled based on the data distributed across multiple parties, and the data from each party cannot be pooled together or provided to other parties. Federated transfer learning was used to address the problem of insufficient data volume and the low number of sample labels. As shown in Table 3, there are large differences in user entities and associated data characteristics between participant A and participant B. The exchange and integration of data between different enterprises is very difficult, but the transfer of federal learning provides a good plan to solve such problems. Nowadays, the heterogeneity problem is still a key challenge in the federated environment. There is a small intersection of dynamic multi-source heterogeneous data in the sample dimension and feature dimension. The federated transfer learning framework under the heterogeneous network ensures the privacy security and the model accuracy and training efficiency of federated learning.

Table 3: Federal Transfer Learning

participant A	ID	X1	X2	Y				
	I1							
	I2							
					participant B	ID	X3	X4	Y
						I3			
						I4			

2.3 Related technologies

2.3.1 Distributed deep learning

The training of a statistical learning model requires a lot of data and is a complex process. With the progress of the model training, the training data of the input model will gradually increase, and the number of parameters will also increase, leading to the whole machine learning process consuming a large amount of resources, such as hardware storage, processing resources, etc. Therefore, people propose distributed learning to realize parallel data and parallel model. Distributed learning uses multiple computational nodes for statistical model training and aims to

improve model training performance and protect privacy. Distributed machine learning uses distributed resources for training to solve the problem that large-scale data training cannot be handled on a single machine. Distributed machine learning uses the parallel data, the parallel model and the parallel pipeline to parallelize the training process, which significantly accelerates the training speed of machine learning and reduces the training time. Among them, deep learning, as one of the advanced methods of machine learning, can learn the internal laws and representation levels of sample data from a large number of complex data, but in some areas that cannot provide enough labeled training data sets, deep learning performance will be greatly reduced. As a mainstream solution, distributed deep learning uses multiple data sources for model training and benefits from large and diverse data sets. However, most of the entities involved in learning are often distributed in various fields, and they are unwilling to contribute their private data for free due to industry rules, legal policies, and privacy issues. Therefore, when distributed deep learning processes privacy data in sensitive fields, it should not only pay attention to model availability, but also pay attention to the privacy security of sensitive data. In FL, as a distributed machine learning method, the distributed client collaboratively trains the machine learning model, while keeping the original data retained locally, with no need to concentrate on a single service node, which not only makes full use of the distributed resources, but also provides privacy security guarantee for the private data of the distributed client.

2.3.2 Convolutional neural network

As a deep learning model, convolutional neural network is dedicated to processing data with network structure, such as images and videos. Convolutional neural network technology plays an important role in image recognition and computational vision tasks. Typical convolutional neural network structures include multiple alternating convolutional and pooling layers, and ultimately fully connected layers. Common models include LeNet, AlexNet, VGG, GoogLeNet (Inception), ResNet, etc., which vary according to the network depth, number of parameters, and task complexity. The convolution operation in convolutional neural network refers to the operation of multiplying a mobile convolution kernel element by element and then adding it. The structure of convolutional neural network is shown in Figure 1, which is mainly divided into four layers, namely input layer, convolution layer, pooling layer and full connection layer. The specific working principle of each layer is as follows.

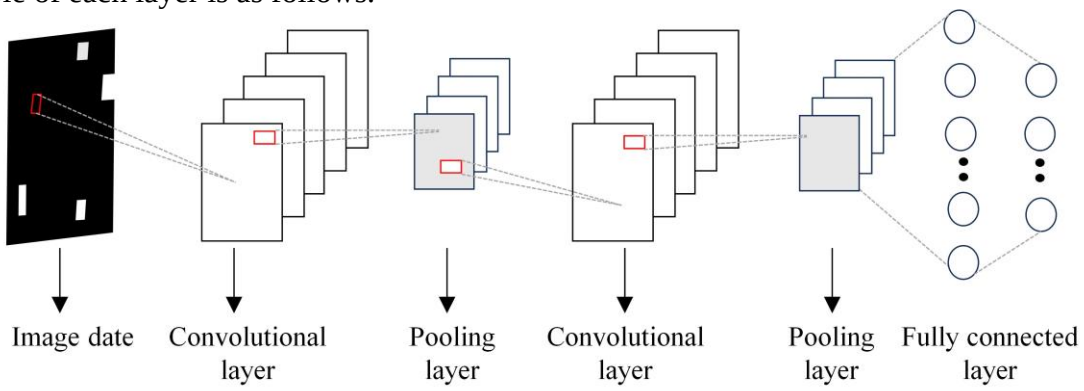

Figure 1 A convolutional neural network structure diagram

(1) The input layer is used to receive data with a network structure, such as images and videos. Compared with the traditional typical artificial neural network transforming the matrix data into a one-dimensional data form, the convolutional neural network retains the spatial correlation of the image data. The image is usually composed of three color channels (red, blue, and green)

(2) The convolutional layer is the core component of convolutional neural networks, and features are extracted from the input image by the convolution operation. The convolution operation slides over the input image through a filter (also called the convolution kernel) and calculates the point product between the filter and the image area each time to generate a feature map (also known as the feature map). These feature maps capture the local features in the image, such as edges, textures, etc.

(3) The pooling layer is used to reduce the spatial dimension of the feature graphs, while retaining the most important feature information. Common pooling operations include maximum pooling (Max Pooling) and average pooling (Average Pooling), which output the maximum or average in the pooling window, respectively.

(4) At the top of the convolutional neural network, one or more fully connected layers are usually connected to use the features extracted by the convolution layer for classification or regression prediction. The output of the fully connected layer is converted into the final prediction result via the softmax function (for the classification problem) or the linear activation function (for the regression problem).

3. Handwritten image recognition based on federal learning

3.1 Network model

In order to solve the huge communication overhead caused by the use of additional security solutions in deep learning, the difficulty to integrate and utilize distributed heterogeneous data, and the difficulty to guarantee the privacy security in the process of model sharing. This paper designs an intelligent recognition system based on federated learning based on the distributed privacy protection learning method of federated learning. The system is mainly composed of a four-layer network architecture, as shown in Figure 2, with the bottom-up terminal layer, edge service layer, network layer and cloud service layer respectively. First, image data collected at the terminal layer is uploaded to the edge service layer. Next, multiple edge computing nodes in the edge service layer perform one or more rounds of local model training on the collected image data. Then, the edge computing nodes transfer the trained local model parameters to the cloud service layer through the secure network layer channel. After that, the cloud service layer performs global model aggregation on all the local model parameters collected and determines whether the global model accuracy has reached the set threshold. If it has, then the training is stopped. Otherwise, the cloud service layer, serving as the aggregation center for federated learning, sends the global model of this round to the edge computing nodes through the network layer, and the model training continues. The main functions of each layer are introduced as follows:

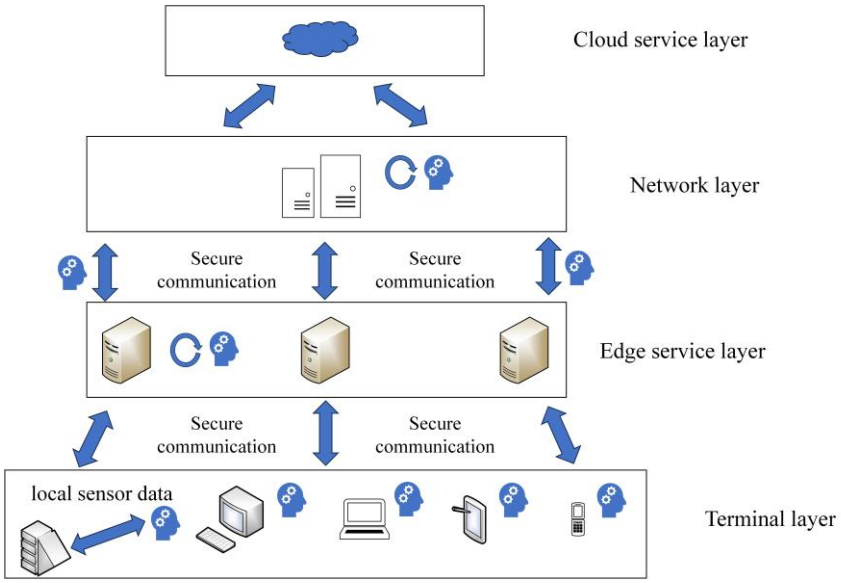

Figure 2 Edge-assisted Internet of Things intelligent identification system based on federated learning

(1) Terminal layer: The terminal layer includes various Internet of Things devices, sensors, smart phones and other intelligent terminal devices. These devices were responsible for collecting and generating data and performing initial processing. For example, a temperature sensor can read the temperature data directly, and a smart camera can capture video images. At the same time, the terminal layer can preprocess and filter the raw data to reduce redundant information and reduce the amount of data transmission. For example, a video surveillance camera can compress the video data locally, or upload only clips involving abnormal activity, rather than the entire video stream. Finally, the terminal layer can provide preliminary security measures, such as data encryption, authentication, etc., to ensure the security and privacy protection of data during transmission and storage. For example, a medical device can encrypt patient data locally and then upload it to the edge or to the cloud.

(2) Edge service layer: The edge service layer is located between the terminal layer and the network layer. It is composed of several edge computing nodes that are responsible for coordinating and managing the local training tasks of various edge terminal devices. The computing nodes in the edge service layer can make local aggregated model updates to the local data of the collected edge terminal devices, conduct local model optimization, and send the aggregated local model updates to the cloud service layer. At the same time, because the edge service layer is close to the data source of terminal devices, real-time response with low latency can be realized, and preliminary security measures, such as authentication, access control, data encryption, etc., to ensure the privacy and security of data transmission and storage processes.

(3) the network layer: network layer is responsible for the edge terminal equipment, edge server and cloud server data transmission and model update, mainly including from the edge computing node and high performance terminal devices to upload local model update to the cloud service layer perform global model aggregation, cloud server to the edge computing node or device issued global model, etc. The network layer uses a secure transmission protocol to encrypt the transmitted model parameters to prevent data eavesdropping and tampering. At the same time, the network layer also dynamically adjusts the transmission frequency and the data quantity, balances the timeliness of the model update and the network load, and manages and optimizes the use of the bandwidth.

(4) Cloud service layer: As the central coordination node of the whole system, the cloud service layer is responsible for managing and coordinating various tasks in the federated learning process.

The cloud service layer is responsible for initializing the global model, sending the initial model parameters to all participating edge servers and edge devices. This is the starting step of federated learning, ensuring that all devices start training from the same model. At the same time, the cloud service layer receives aggregated model updates (such as gradient or model parameters) from each edge server to conduct global aggregation. This usually involves weighted averaging, incorporating the contributions of each edge device into an updated global model. Finally, the cloud service layer distributes the aggregated global model parameters back to the edge server, and then the edge server is further distributed to each edge device for the next round of local training. This process was repeated for each round of the federated learning iteration until the model converged.

3.1 Training of intelligent recognition model based on federated learning

In order to solve the traditional privacy security scheme of huge communication overhead, ensure the privacy in the process of model sharing problems, in this paper, the system using federal learning design intelligent recognition system, using centralized federal learning architecture to solve traditional deep learning needs to upload the original local data to the central server training problem, avoid the direct leakage of raw data. At the same time, the federated learning architecture allows multiple data sources to cooperate, alleviating the problem of difficult to use heterogeneous data. The training process of the intelligent recognition model based on federated learning is as follows:

(1) Global initialization: The cloud service center initializes the global model, transmits through the security channel of the network layer, to all edge computing servers, and then by the edge server to each edge terminal device that can independently conduct local model training.

(2) Local model training: Each high-performance terminal device will use the local data set for the global model training, and calculate the model updates. For some edge terminal devices with limited performance, it is necessary to upload the local data to the nearest edge computing server to perform one or more rounds of local model training.

(3) Local model update and upload: high-performance edge terminal devices and edge computing servers will use differential privacy technology to add noise to the local model update, and then upload it to the cloud center server for global model aggregation;

(4) Global model aggregation: The cloud center server updates the global aggregation of the received encrypted local model, and updates the global model. In this paper, the system adopts the federated average aggregation algorithm to aggregate the global model;

(5) Model distribution: The cloud center server judges that the current global model will send the updated global model to the edge server, and then be distributed from the edge server to the edge device, and enter the next round of iteration until the model converges.

4. Experimental results and analysis

In order to verify the effectiveness of the intelligent recognition system based on federated learning proposed in this paper, this section has set up simulation experiments to verify the effectiveness of the proposed system on the MNIST dataset, to discuss and evaluate the model accuracy under different privacy budgets, and to design the comparative experiment to verify the advanced nature of the proposed system.

4.1 Experimental environment setting

The experimental environment configuration is as follows: Intel (R) Core (TM) i7-9700 CPU @ 3.00GHz, 32GB memory, Windows11 operating system. The experiment constructed the FL

framework using TensorFlow 2.0 and used CNN as the training model.

4.2 Data set

This paper uses the MNIST handwritten image data set generated from downsampling in real environment to verify the effectiveness of the proposed system. The MNIST dataset is an important benchmark dataset in machine learning and computer vision, widely used for handwritten digit recognition tasks. The MNIST dataset consists of 60000 training samples and 10000 test samples. Each sample is a 28x28 pixel grayscale image representing a handwritten number between 0 and 9. The data set was provided by the National Institute of Standards and Technology (NIST). Images in the MNIST dataset were all gray sizes with 28 * 28 pixels, each with values between 0 and 25, indicating different gray levels. Each image has a corresponding label, indicating the number (0-9) in the image, and the label is an integer, with 10 categories. In order to construct a real distributed learning scenario, 10 client devices were set up in this section, and the whole MNIST dataset was randomly divided into 10 equal parts, and each subset was sent to one client device for local model training.

4.3 Hyper-parameter setting and evaluation indicators

In this section, the global number of iterations is set as 100, the number of local training rounds is 3, the learning rate is 0.005, the local training batch size of the MNIST dataset is 10, and the stochastic gradient optimization algorithm of descent was used for the local training.

In this section, the accuracy rate and accuracy rate is used to evaluate the performance of the model by calculating formulas such as Equations (1) and (2):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$precision = \frac{TP}{TP + FP} \tag{2}$$

Specifically, TP was true positive, TN was true negative, FP was false positive, and FN was false negative.

4.4 Comparison of the experimental results

In order to verify the effectiveness of the proposed federated learning-based intelligent recognition system, the model accuracy of the experiment in this section compares the proposed system with scheme 1 [15] and Scheme 2 [16] on the MNIST dataset. Both schemes 1 and 2 enable deep learning to train the image recognition model without adding additional privacy security scheme. Figure 3 shows that the model accuracy of the proposed system scheme is as high as 98.7%, which is higher than that of comparing scheme 1 and Scheme 2, which proves the effectiveness of the proposed federated learning-based intelligent recognition system.

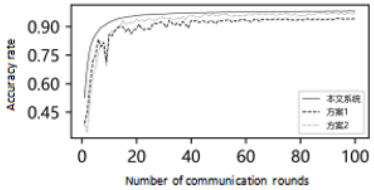

Figure 3 Comparison of model accuracy for the different schemes

Table 4 statistics the experiment in this section compares different schemes on the MNIST data set model accuracy and precision, as shown in table 4, the proposed system after added additional privacy security mechanism still achieved the highest model accuracy, 1.6% higher than contrast scheme 2, and contrast scheme 1 and scheme 2 did not add additional privacy security mechanism, local data uploaded to the central server to perform training task process there are serious data security problems and privacy leakage risk. Malicious enemies can attack the central server to tamper with the model parameters, destroy the model accuracy, or steal the original data set information of local users during training. Since the training of deep learning model requires a large amount of user data, once the central server is attacked by malicious enemies or collude with it, it may lead to large-scale private data leakage or damage to the accuracy of the model, causing huge losses to users. However, the system mentioned in this paper only needs to upload the local model to update to the cloud center server for global model training, and the original data of users is retained locally, which effectively alleviates the risk of users' local privacy data leakage. At the same time, the proposed system in this paper uses differential privacy technology to noise the local model update to ensure the privacy security in the transmission process and global model aggregation. Even if the malicious enemy obtains the local model update after the noise of the local model update, the user's original privacy data cannot be accurately inferred. In addition, due to the different data distribution owned by different clients, the accuracy of the model trained with a potentially huge sample size is not high, and some clients are not willing to actively contribute their own local data parameter model training due to privacy security issues. The proposed system in this paper uses the federated learning training paradigm to collaboratively train the global model in a privacy-protected way, which effectively breaks the "data island" and can cooperatively train multi-party data.

Table 4 Comparasting results of accuracy and accuracy for different schemes

data set	Contrast the scheme	precision (%)	accuracy (%)
MNIST	This paper scheme	98.7	99.2
	Scheme 1	93.7	95.5
	Scheme 2	97.1	98.9

The intelligent recognition system based on federated learning uses the local model parameters to upload the local model parameters before the edge computing server. However, the local differential privacy setting may have some influence on the local model parameters and the accuracy of the global model. Therefore, this section experiments on the statistical verification and comparison of the accuracy of federated learning global models under different privacy budgets, and presents a process for the discussion of privacy budget parameters. Figures 4,5,6 and 7 respectively count the impact of different privacy budget settings on the accuracy of the federated learning global model.

The experimental results in Figure 4 shows the accuracy of the federated learning global model without adding any privacy noise. The traditional federated learning model without adding privacy policy is stable by about 98% in the 100th round of communication, and the model performance is better. Among them, figures 5,6 and 7 respectively count the accuracy of the federated learning global model when adding the privacy noise-protected client local model update. It can be seen that when the local differential privacy technology is used to protect the local model parameters, the accuracy of the global model will be affected and the accuracy will decrease. Moreover, the convergence rate of the model are different under different privacy budget Settings.

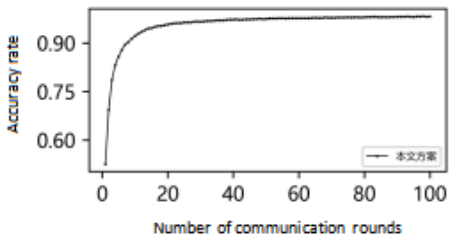

Figure 4 Global model accuracy without adding privacy noise

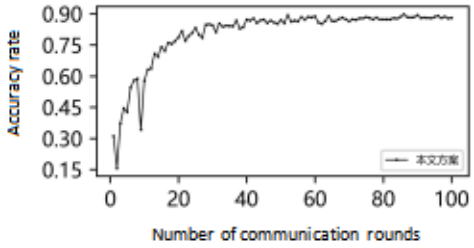

Figure 5 Global Model Accuracy (privacy budget set to 2.5)

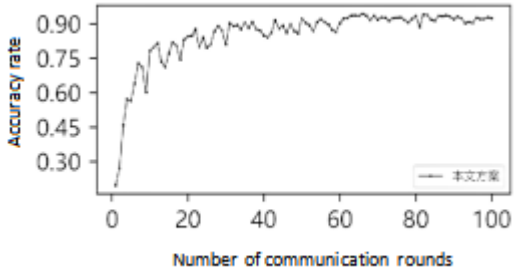

Figure 6 Global Model Accuracy (privacy budget set to 3.0)

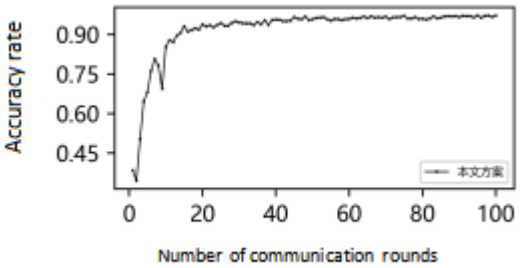

Figure 7 Global Model Accuracy (privacy budget set to 3.5)

Figure 5 shows a maximum global model accuracy of 89.2% when the privacy budget was set to 2.5. At the early stage of training, due to the addition of Laplace noise to protect the privacy of local model parameters, the accuracy of the model fluctuates at this time and is not stable. With the increase of the number of training rounds, the accuracy of the model gradually stabilized, reaching the highest accuracy of 89.2% in 100 rounds of training. Figure 6 shows the global model accuracy when the privacy budget is set to 3.0, and the model accuracy gradually stabilizes after 100 rounds of training, up to 91.8%. Figure 7 shows the global model accuracy with the privacy budget set to 3.5, with the highest model accuracy of 94.2% after 100 rounds of training.

Table 5 statistics the different privacy budget set of the federal learning global model accuracy, can be seen from table 5, with the increase of privacy budget, global model accuracy is improving, this is because the privacy budget and noise ratio is different, the higher the privacy budget, add the less noise, the less is the influence of the model accuracy, so can in much influence model accuracy can play the role of privacy data protection. At the same time, if the noise is added too little, then the effect of privacy protection is not good.

Table 5 Comparison results of model accuracy for different privacy budget settings

Privacy budget	precision (%)
No noise added	98.2
2.5	89.2
3.0	91.8
3.5	94.2

5. Conclusion

This paper proposes a kind of federal learning based on federal learning edge auxiliary Internet of things intelligent recognition system, based on the terminal layer, edge service layer, network

layer and cloud center service layer four layer architecture design based on federal learning intelligent recognition system, in the way of privacy protection auxiliary Internet devices for safe and intelligent object recognition. Secondly, this paper experimentally verifies and analyzes the effectiveness and accuracy of the intelligent recognition system based on federated learning, and the results show that the proposed system can achieve high model accuracy through privacy protection. However, the use of differential privacy technology still has a balance between privacy and effectiveness, and subsequent research will further improve the use of differential privacy technology and reduce the impact on model accuracy while protecting model privacy.

References

[1] Bayoudh K. A survey of multimodal hybrid deep learning for computer vision: Architectures, applications, trends, and challenges[J]. Information Fusion, 2023: 102217.

[2] Saini M, Susan S. Tackling class imbalance in computer vision: a contemporary review[J]. Artificial Intelligence Review, 2023, 56(Suppl 1): 1279-1335.

[3] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.

[4] Menghani G. Efficient deep learning: A survey on making deep learning models smaller, faster, and better[J]. ACM Computing Surveys, 2023, 55(12): 1-37.

[5] Dou Hui, Zhang Lingming, Han Feng, et al. Review of interpretability studies of convolutional neural networks [J]. Journal of Software, 2024,35 (01): 159-184.

[6] Talaei Khoei T, Ould Slimane H, Kaabouch N. Deep learning: Systematic review, models, challenges, and research directions[J]. Neural Computing and Applications, 2023, 35(31): 23103-23124.

[7] Mcmahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Artificial intelligence and statistics. Ft Lauderdale, USA: PMLR, 2017: 1273-1282.

[8] Xiao Xiong, Tang Zhuo, Xiao Bin, et al. Review of Privacy Protection and Security Defense Research in Federal Learning [J]. Journal of Computer Science, 2023,46 (05): 1019-1044.

[9] KhoKhar F A, Shah J H, Khan M A, et al. A review on federated learning towards image processing[J]. Computers and Electrical Engineering, 2022, 99: 107818.

[10]Chaddad A, Wu Y, Desrosiers C. Federated learning for healthcare applications[J]. IEEE Internet of Things Journal, 2023, 11(5): 7339-7358.

[11] Muhammed D, Ahvar E, Ahvar S, et al. Artificial Intelligence of Things (AIoT) for smart agriculture: A review of architectures, technologies and solutions[J]. Journal of Network and Computer Applications, 2024: 103905.

[12] Pandya S, Srivastava G, Jhaveri R, et al. Federated learning for smart cities: A comprehensive survey[J]. Sustainable Energy Technologies and Assessments, 2023, 55: 102987.

[13] Zhou J, Lu Q, Dai W, et al. Guest editorial: Federated learning for industrial IoT in industry 4.0[J]. IEEE Transactions on Industrial Informatics, 2021, 17(12): 8438-8441.

[14] Li H, Ge L, Tian L. Survey: federated learning data security and privacy-preserving in edge-Internet of Things[J]. Artificial Intelligence Review, 2024, 57(5): 130.

[15] Xu W, Li W, Wang L. Research on Image Recognition Methods Based on Deep Learning[J]. Applied Mathematics and Nonlinear Sciences, 9(1): 1-14.

[16] Zhang H, Li J, Liu S, et al. A Human Body Infrared Image Recognition Approach via DCA-Net Deep Learning Models[J]. International Journal on Artificial Intelligence Tools, 2023, 32(05): 2360004.