Collaborative Deep Learning for Medical Image Analysis with Differential Privacy

Danni Yuan*, Xiaoyan Zhu* , Mingkui Wei†, Jianfeng Ma*
*State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China
†College of Science and Engineering Technology, Sam Houston State University, Texas 77340, USA
Email: dnyuan@stu.xidian.edu.cn, xyzhu@mail.xidian.edu.cn, mwei@shsu.edu, jfma@mail.xidian.edu.cn

***Abstract*—Deep learning algorithms, especially convolution neural networks, have attracted huge attention in the field of medical image analysis. A hospital could train a neural network to detect disease based on medical images possessing. However, the number of medical images would affect the results of training. If medical images of all hospitals are collected together, there's a risk of privacy leakage. In this paper, we apply collaborative deep learning to medical image analysis, which could help to improve the training effect. Besides, we also exploit differential privacy, the analytic Gaussian Mechanism, to prevent the leakage of information about medical images. We experiment on the Chest X-ray Images (Pneumonia) dataset. Results show that the analytic Gaussian Mechanism can protect the privacy of medical images effectively, while the influence on the results of training is small. The accuracy can be improved about 19% via collaborative deep learning and can still remain about 18% even when the analytic Gaussian Mechanism was used.**

***Index Terms*—collaborative deep learning, medical image analysis, differential privacy**

I. Introduction

In recent years, deep learning has been used in many fields of engineering, ranging from text processing [1], speech recognition [2] to computer vision [3]. Especially, the ImageNet competition in December 2012 successfully brought Convolution Neural Networks (CNNs), called AlexNet, to the public, which attracted many researchers' attention [3]. This is the first time that a neural network uses ReLU as activation function and trains a model through Graphics Processing Units (GPUs). After that, deep convolution networks have become a preferential choice in computer vision.

The medical image analysis community has also paid attention to these developments. According to the survey about deep learning in medical image analysis [4], the number of papers grew explosively since 2015. There are various networks applied to the medical field, such as Deep Belief Networks (DBNs) [5], Stacked Auto-Encoders (SAEs) [6], Restricted Boltzmann Machines (RBMs) [7], Recurrent Neural Networks (RNNs) [8] and CNNs [9]. Comparing with other methods to analyze medical images, researchers prefer to choose CNNs whose applications are ranging from pathology, brain, cardiac to lung and abdomen.

However, a hospital using deep learning in medical image analysis probably encounters a problem with the small number of training dataset. Tajbakhsh et al. [10] pointed out that it might be difficult in the medical domain to require a large amount of labeled training data because of expensive expert annotation and the scarce diseases (e.g.,lesions). The whole labeled medical training data gathered worldwide is relatively less than data gathered in other fields, leaving alone the labeled medical data gathered in merely one single hospital. Indeed, this is a common phenomenon in the field of medical image analysis.

Since the labeled medical dataset gathered in one hospital might be small, which leads to less accurate outcome of the training model, collaborative deep learning suits the situation perfectly. In other words, the labeled medical data can be trained in a local hospital, then the weight of local neural network can be uploaded to a parameter server and can be shared with other hospitals. This could improve the accuracy of each neural network in each hospital just like the dataset is expanded.

Furthermore, the process of sharing parameters has a potential risk of medical data leakage. Although only the parameters of local neural network is uploaded, the original data can be recovered according to it. Shokri et al. [11] mentioned that a fraction of the neural-network parameters would reveal some information about training datasets indirectly during training process. And Aono et al. [12] introduced how a small portion of gradients of neural network may reveal information about local data in details. Hitaj et al. [13] mentioned that the information of dataset could be recovered by Generative Adversarial Network (GAN). However, privacy plays a key role when dealing with medical data. No one wants his/her medical data under the risk of leakage. Health Insurance Portability and Accountability Act (HIPAA) also mandates on privacy protection of patients' medical records.

Aiming to protect privacy during the process of collaborative deep learning, several methods have been introduced. Shokri et al. [11] proposed a scheme named Selective Stochastic Gradient Descent (SSGD) in which the participants in the distributed system could select a few fraction of gradients to upload, according to the parameter exchange protocol. Besides, they also preserved privacy by using differential privacy, which added the Laplace noise to the selective gradients. In order to measure the total privacy loss, Abadi et al. [14]

This work was supported in part by projects of Nature Science Foundation of China under Grant 61772406 and Grant U1636209, in part by the project of Fundamental Research Funds for the Central Universities under Grant JB180110. *(Corresponding author: Xiaoyan Zhu)*

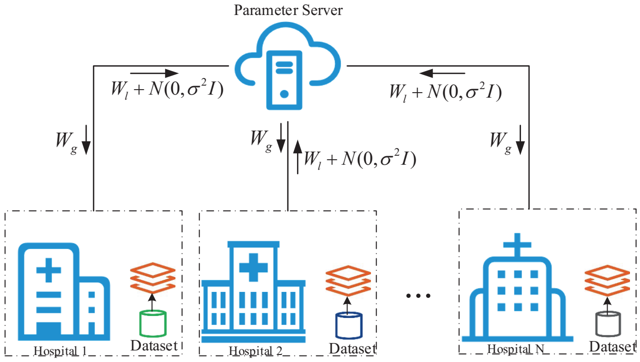


Fig. 1. Collaborative deep learning for medical image analysis with differential privacy.

proposed the moments accountant which could additively accumulate the log of the moments of the privacy loss at each training step.

In this paper, to solve the problems of small dataset and privacy leakage, we propose a collaborative deep learning method for medical image analysis with differential privacy. The highlights of our main contributions are summarized as follows:

- We apply collaborative deep learning to the field of medical image analysis to solve the problem of small dataset in this field. And in the experiment, we assume there exists a parameter server who has responsibility to store and distribute sharing parameters. Considering the computer in a hospital might be a device with low computation power, we adopt AlexNet [3] as our neural network in the collaborative deep learning.
- We use differential privacy to protect the privacy of medical images, considering the leakage caused by sharing parameters. In the experiment, we add Gaussian noise to sharing parameters, which satisfies the definition of differential privacy. To acquire the smaller standard deviation of Gaussian noise while the privacy cost and the additive term are remained, we use the analysis Gaussian Mechanism (aGM) proposed by Balle et al. [15] rather than the classical Gaussian Mechanism (cGM) proposed by Dwork et al. [16].
- We achieve collaborative deep learning on the Chest X-ray Images (Pneumonia) dataset and add the analytic Gaussian mechanism to sharing parameters. Results show that collaborative deep learning could improve the accuracy about 19%. The accuracy is still as high as 17% after Gaussian noise was added.

The rest of this paper is organized as follows. In Section II, we briefly discuss the related work about collaborative deep learning and differential privacy. The details of our scheme and algorithms we used are introduced in Section III. And Section IV describes our experimental results. Finally, we conclude the paper in Section V.

II. PRELIMINARIES

In this section, we introduce convolutional neural networks and the theorem of differential privacy briefly. In addition, we also introduce former work combining differential privacy with deep learning.

A. *Convolutional Neural Networks*

CNNs have a great performance in the fields of recognizing patterns from images through feature extraction and classification. A typical architecture of CNNs consists of four types of layers: convolutional, activation, pooling and fully-connected (or dense) layers.

Convolutional layers are used to detect certain local features of input images. In this process, each neuron of a convolutional layer is only connected to a small area of the input image. Besides, to enable the search for a same feature throughout the images, the weights are shared between neurons in the convolutional layers, each set of shared weights is a kernel. And to obtain more complex properties of the image, there is usually the non-linear activation layer behind convolution layers. Pooling layers are often used to subsample previous layers by reducing the outputs of neuron clusters into a single neuron, which may compute a max or an average. Max pooling uses the maximum value and average pooling uses the average value. At last, one or more fully-connected (or dense) layers are used to produce the classification results, each of which is followed by an activation layer.

B. *Differential Privacy*

Differential privacy is a definition of privacy, not an algorithm, addressing the problem that the data of a population could be used while the individual information is protected. The mechanisms of differential privacy mainly include the Laplace mechanism [17], the exponential mechanism [18] and the Gaussian mechanism [19]. In our experiments, we use the Gaussian mechanism, which perturbs each elements with the noise drawn from the Gaussian distribution. The definition of differential privacy is as follows:

Definition 1: (ϵ, δ)-Differential Privacy [20]. A randomized mechanism $\mathcal{M} : \mathcal{D} \to \mathcal{R}$ with domain $\mathcal{D}$ and range $\mathcal{R}$ satisfies (ϵ, δ)-differential privacy if for any two neighboring inputs $d, d' \in \mathcal{D}$ and for any subset of outputs $\mathcal{S} \subseteq \mathcal{R}$ it holds that

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

The real-valued function $f : \mathcal{D} \to \mathcal{R}$ would be perturbed through additive noise with differential privacy mechanism, and the sensitivity S_f of it is defined as $sup_{d \simeq d'} \|f(d) - f(d')\|_2$, where $\| \|_2$ means the L_2 norm of $\|f(d) - f(d')\|$.

Theorem 2: Gaussian Mechanism [16]. The Gaussian mechanism is defined as follow:

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

where the $\mathcal{N}(0, S_f^2 \cdot \sigma^2)$ represents the Gaussian distribution which the mean and variance is 0 and $S_f^2 \cdot \sigma^2$, respectively.

Let $\sigma = \sqrt{2log(1.25/\delta)}/\epsilon$, the mechanism could satisfy the (ϵ, δ)-differential privacy for any $\epsilon, \delta \in (0, 1)$. Hence, the Gaussian mechanism provide an approach to calibrate a zero mean isotropic Gaussian perturbation $Z \sim \mathcal{N}(0, S_f^2 \cdot \sigma^2)$ to the L_2 sensitivity of f.

C. Differential Privacy in Deep Learning

It's the seminal work [11] that noticed the problem of privacy in the deep learning network and applied differential privacy to fix it. To avoid leaking the privacy, they add noise to uploaded parameters which satisfied differential privacy mechanism. One drawback of this scheme is that the accuracy of network is a bit low. In order to advance the scheme, Abadi et al. [14] come up with the moments accountant which approve their algorithm satisfied $(O(q\epsilon\sqrt{T}), \delta)$ - differential privacy with $\sigma \geq c\frac{q\sqrt{T\log(1/\delta)}}{\epsilon}$, where c is a constant, q is the probability of picking each example $q = \frac{L}{N}$ and N is the size of the input dataset. In their algorithm, gradients are bounded by clipping each gradient in L_2 norm with the predefined threshold C.

$$\overline{g}_t \leftarrow g_t(x_i)/max(1, \frac{\|g_t(x_i)\|_2}{C}) \tag{3}$$

$$\tilde{g}_t \leftarrow \frac{1}{L}\left(\sum_i \overline{g}_t(x_i) + \mathcal{N}(0, \sigma^2 C^2 I)\right) \tag{4}$$

$$w_{t+1} \leftarrow w_t - \eta\tilde{g}_t \tag{5}$$

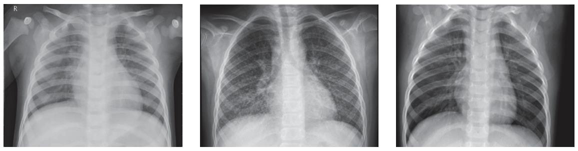
At last, they demonstrate that they improved the quality of the model and preserved the privacy of sharing parameters by training a deep neural network with differential privacy under a modest privacy budget.

III. OUR APPROACH

In this section, we introduce the main components of our approach toward collaborative deep learning for medical image analysis and privacy preserving of training dataset via differential privacy: the dataset we used in the study, the architecture of proposed CNN and the method of applying differential privacy.

A. Data

In this part, we use the Chest X-Ray Images (Pneumonia) dataset which was collected by Kermany et al. [21]. The dataset was collected from a total of 5,856 patients and divided into a training set and a test set. The training set contains 5,232 chest X-ray images (JPEG) from children, including 3,883 characterized as depicting pneumonia (2,538 bacterial and 1,345 viral) and 1,349 normal. The test set collected from patients contains 624 chest X-ray images (JPEG) with 234 normal images and 390 pneumonia images (242 bacterial and 148 viral). Fig. 2 shows three kinds of chest images, bacterial (left), viral (middle) and normal (right). Bacterial pneumonia typically shows a focal lobar consolidation. Viral pneumonia exhibits a more diffuse "interstitial" pattern in both lungs. But normal chest X-ray images manifest with clear lungs without any areas of abnormal opacification in the image.



a) Bacterial Pneumonia b) Viral Pneumonia c) Normal

Fig. 2. Examples of chest X-ray images in the dataset

B. Neural-Network Architecture

In this part, we introduce our neural-network architecture. Considering the detection of pneumonia in chest X-ray images, we decided to choose convolution neural networks which have good performance in image classification. We adopt the AlexNet model as the backbone of the network. And our experiment is divided into several parts as follows:

- To verify the effect of collaborative deep learning in the field of medical image analysis, we compare two different scenarios one of which is sharing parameters between those hospitals who take part in collaborative deep learning, the other is training a neural network individually with the local dataset without sharing parameters.
- We explore the difference of accuracy between collaborative deep learning on AlexNet [3] and larger-scale deep neural networks but without sharing parameters such as VGG-16 [22] and DenseNet-121 [23]. If the accuracy is similar or even better, that means hospitals can get better results without updating their hardware.

Then we introduce the architecture of the AlexNet. In fact, we also have made some changes in the AlexNet such as activation functions and fully connected layers.

The AlexNet contains eight layers with weights, five convolutional and three fully-connected layers. The output of the last fully-connected layer is the input of a 1000-way softmax which produces a distribution over the 1000 class labels. The first convolutional layer uses 96 kernels of size $11 \times 11 \times 3$ to filter the $227 \times 227 \times 3$ input image. The second convolutional layer uses 256 kernels of size $5 \times 5 \times 48$ to filter the output of first convolutional layer. And the third convolutional layer has 384 kernels of size $3 \times 3 \times 256$ connected to the output of second layer. The fourth convolutional layer has 384 kernels of size $3 \times 3 \times 192$. The fifth convolutional layer has 256 kernels of size $3 \times 3 \times 192$. Finally, each fully-connected layer has 4096 neurons.

It is known that the activation function of the AlexNet is the ReLU function $f(x) = max(0, x)$. Comparing with the classical sigmoid, the use of the ReLU function has been proven to speed up the training process for many times. But ReLU completely suppresses negative values, which would cause the "dying ReLU" problem. Driven by this consideration, we decide to use a variant of ReLU, the LeakyReLU.

Different from ReLU, LeakyReLU assigns a non-zero slope (6).

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha x, & \text{else} \end{cases} \tag{6}$$

where α denotes a manually set coefficient.

In addition, our dataset is 5,856 images belonging to 2 categories (normal and pneumonia), we change the number of neurons of the last layer from 1,000 to 2. To satisfy the requirement of input in the AlexNet, we extract random 224×224 patches from our 944×640 chest X-ray images.

C. Collaborative Deep Learning

As to the collaborative deep learning, it's important to manage the way of sharing parameters between hospitals. As introduced by Shokri et al. [11], there exists two approaches to upload parameters in collaborative deep learning, round robin, random order and asynchronous. In our system, we mainly consider the round robin exchange protocol.

In the round robin exchange protocol, a weight matrix and learning rate are initialized by the global server. Then, the global server distributes the initial weight to the first hospital and receives the parameters uploaded by the first hospital. After receiving parameters, the global server sends it to the next hospital. As for all hospitals, they upload and download parameters following a fixed order. This order is unchanged during the entire training process.

D. Privacy Preserving

Considering the sharing parameters would leak the privacy of training samples [11]–[13], we use differential privacy in our scheme. To acquire the smaller σ while the privacy loss remains low, we decide to apply the Gaussian Mechanism which add the Gaussian noise to sharing parameters. But, it is known that adding much noise to sharing parameters would destroy the utility of sharing parameters and influence the accuracy of neural network. However, the small variance of noise could not satisfy the demand for preserving privacy. This is exactly the game between the utility and privacy-preserving. Therefore, we prefer to find a way that achieves the balance between utility and privacy-preserving in collaborative deep learning system.

In order to reduce the variance of noise and improve the classification accuracy, we adopt the aGM which could achieve the smaller noise in the same level of privacy cost compared with the cGM. The theorem of the analytic Gaussian mechanism is introduced as follow:

Theorem 3: Analytic Gaussian Mechanism [15]. For any $\epsilon \geq 0$ and $\delta \in [0, 1]$, the Gaussian mechanism $\mathcal{M}(d) = f(d) + Z$ is (ϵ, δ)-differential privacy if and only if

$$\Phi\left(\frac{S_f}{2\sigma} - \frac{\epsilon\sigma}{S_f}\right) - e^{\epsilon}\Phi\left(-\frac{S_f}{2\sigma} - \frac{\epsilon\sigma}{S_f}\right) \leq \delta \tag{7}$$

where $\Phi()$ represent the CDF of the standard Gaussian distribution $\Phi(t) = \mathbb{P}[\mathcal{N}(0,1) \leq t] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-y^2/2} dy$.

Our noise scale σ is calculated from specific ϵ and δ according to Theorem 9 in [15]. The complete training procedure for a participant is described in Algorithm 1 which explains the whole training procedure of all hospitals in the collaborative deep learning. We assume that the local training procedure is safe and reliable. We add the Gaussian noise to the sharing parameters. The global server receives and preserves the weights which will be shared to other hospitals. Even though the global server and other hospitals get the sharing parameters, they couldn't know the precise values. In this way, the privacy of hospitals could be protected and the classification accuracy of neural network could be improved.

Algorithm 1 Collaborative Deep Learning with the analysis Gaussian Mechanism

```
Input: The number of all hospitals $N$, training dataset of all hospitals $D = \{D_1 \ldots, D_N\}$, the initial weight in the global server $W_g^{(0)}$, the standard deviation of Gaussian noise $\sigma$, the training epoch for hospital $T_i$.
for $i \in N$ do
    Download weight $W_g^{(i-1)}$ from the global server.
    for $t \in T_i$ do
        Local training process with the AlexNet
        if $t = T_i$ then
            Add noise $W_g^{(i)} = W_l^{(T_i)} + \mathcal{N}(0, \sigma^2 I)$
        end if
    end for
end for
Output: Upload $W_g^{(i)}$ to the global server
```

IV. Experimental Results

This section mainly focuses on the presentation and discussion of the results. At the beginning, we introduce our experimental implementation. The experiment, about collaborative deep learning, was preformed under a Windows 10 system on a devices with CPU Inter(R) Core(TM) i3-7100 @ 3.90GHz, GPU NVIDIA GeForce GTX 1050, and 8 GB of RAM. However, the experiment, about large-scale neural network like VGG-16 and DenseNet-121, was performed on a device with CPU Inter(R) Core(TM) i7-8750H @ 2.20GHz, GPU NVIDIA GeForce GTX 1060. Besides, all experiments used the PyTorch framework, coded in python and completed with GPU.

A. The Effect of Collaborative Deep Learning

In this subsection, we measure the effect of collaborative deep learning. In this experiment, all hospitals use the neural network, AlexNet introduced before. We design the experiment from two aspects: the accuracy of collaborative deep learning, the performance compared with training a neural network individually by AlexNet, VGG-16 and DenseNet-121.

To analyze the accuracy of collaborative deep learning, we assume that there are four groups which have different number of hospitals and training images. In this experiment, the Group1 has 5 hospitals and each of them has 1000 training

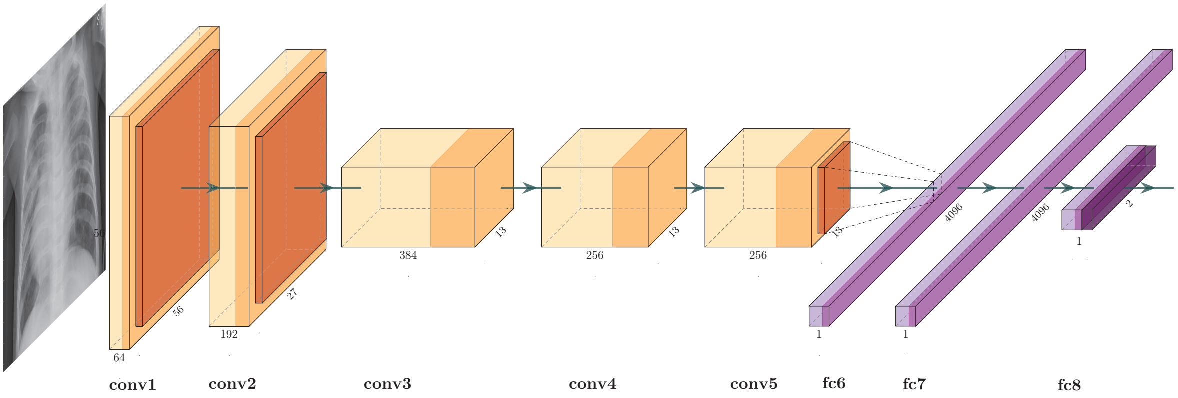


Fig. 3. Architecture of the neural network for the detection of pneumonia with the chest X-ray images

chest X-ray images and 624 test chest X-ray images. The Group2 has 10 hospitals and each of them has 500 training chest X-ray images and 624 test chest X-ray images. The Group3 has 15 hospitals and each of them has 300 training chest X-ray and 624 test chest X-ray images. And the Group4 has 20 hospitals and each of them has 200 training chest X-ray and 624 test chest X-ray images. During the experiment, we set the training epoch of a hospital as 40.

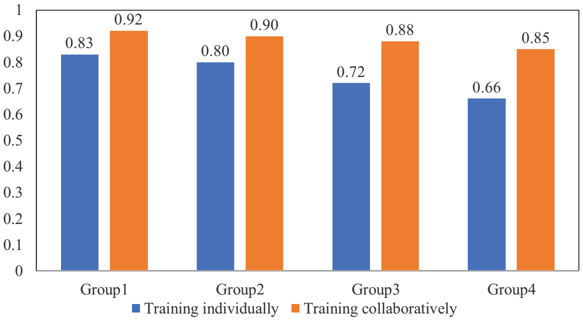


Fig. 4. Comparison of test accuracy between four groups. The Group1 has 5 hospitals and each has 1000 training images, which is referred to as (5, 1000). And the rest of groups are (10, 500), (15, 300), (20, 200), respectively.

The experimental result has been shown in Fig. 4. In Fig. 4, we found that when a hospital trains a neural network individually with the smaller number of training images, the classification accuracy would be lower. This phenomenon is obvious especially in the Group3 and Group4 where the values of test accuracy are only 0.72 and 0.66, respectively. However, when a hospital trains a neural network via collaborative deep learning, the values of test accuracy are all higher than training individually and the highest accuracy can be as high as 0.92.

To explore whether collaborative deep learning with AlexNet could achieve the similar performance of training individually by AlexNet, VGG-16 and DenseNet-121, we assume there's a hospital which has 1000 training chest X-ray images and 624 test chest X-ray images. In the first condition, the hospital trains a neural network individually with AlexNet, VGG-16 and DenseNet-121 without sharing parameters. In the second condition, the hospital trains a neural network by collaborative deep learning with AlexNet. Then we can compare the training performance in these two conditions.

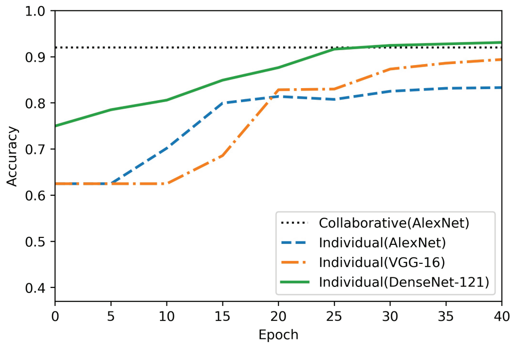


Fig. 5. The training effect of different convolution neural networks.

From Fig. 5., we could find out that in the first condition, the accuracy of DenseNet-121 is higher than the accuracy of VGG-16, while the accuracy of AlexNet remains the lowest. Although DenseNet-121 and VGG-16 have good performance in the classification of the chest X-ray images, they also ask for higher requirements to the hardware. However, in Fig. 5., it shows that a hospital possessing AlexNet could acquire a similar or higher accuracy compared with DenseNet-121 or VGG-16 via collaborative deep learning.

B. *Preserving-Privacy Deep Learning in Medical Images*

In order to protect sharing parameters, we inject Gaussian noise to sharing parameters and make it satisfy the definition of differential privacy. Different from the previous works [11], [14] which adopted the Laplace Mechanism or the classical Gaussian Mechanism, we introduce the analytic Gaussian Mechanism in our scheme. The analytic Gaussian Mechanism could decrease the variance of Gaussian noise while the privacy cost remained. The process of adding the analytic Gaussian Mechanism to sharing parameters was described in the Algorithm 1.

Besides, we explore the relationship between the privacy loss and the accuracy of neural network. we add various Gaussian noise to sharing parameters in the collaborative deep learning. It is known that the standard deviation of the Gaussian noise is computed by the privacy cost ϵ and the additive term δ. Indeed, when ϵ or δ tends to small, the standard deviation of the Gaussian noise would become large. There are three various Gaussian noise in our experiment which are the large one $(0.5, 10^{-5})$, the medium one $(2, 10^{-5})$ and the small one $(8, 10^{-5})$.

TABLE I
ACCURACY FOR FOUR GROUPS WITH DIFFERENTIAL PRIVACY

(ϵ, δ)	Noise size	Group1	Group2	Group3	Group4
None	$\sigma = 0.000$	0.92	0.90	0.88	0.85
$(8, 10^{-5})$	$\sigma = 0.600$	0.90	0.89	0.87	0.84
$(2, 10^{-5})$	$\sigma = 1.993$	0.88	0.87	0.84	0.82
$(0.5, 10^{-5})$	$\sigma = 7.032$	0.87	0.85	0.82	0.79

Table I shows the classification accuracy when apply differential privacy to the sharing parameters. From Table I, we discover that adding the small Gaussian noise, such as $\sigma = 0.600$, to sharing parameters during the training process of collaborative deep learning would not decrease test accuracy of classification too much. And through applying differential privacy, we could achieve the protection of privacy of medical images. However, the large noise size, such as $\sigma = 7.0322$, would affect test accuracy obviously, especially when the hospital possesses the small number of medical images. Although adding Gaussian noise to the sharing parameters would decrease more or less the classification accuracy, the effect of collaborative deep learning with differential privacy is still better than training neural network individually. So, in reality scenario of medical image analysis, it is necessary to find an appropriate noise level to achieve privacy preserving while the accuracy remained.

V. CONCLUSIONS

In this paper, we apply collaborative deep learning and differential privacy to the field of medical image analysis. We test it based on the Chest X-Ray Images (Pneumonia) dataset. Results show that collaborative deep learning increases the classification accuracy effectively, especially when there are relatively fewer training samples. After adding Gaussian noise to sharing parameters, the decrease of accuracy is small and affects little to the results. In the further work, we would like to consider other methods of cryptography, such as secure multiparty computation or homomorphic encryption.

REFERENCES

[1] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, “Learning word vectors for sentiment analysis,” in *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1*. Association for Computational Linguistics, 2011, pp. 142–150.

[2] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath *et al.*, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal processing magazine*, vol. 29, no. 6, pp. 82–97, 2012.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[4] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[5] A. Ortiz, J. Munilla, J. M. Gorriz, and J. Ramirez, “Ensembles of deep learning architectures for the early diagnosis of the alzheimer’s disease,” *International journal of neural systems*, vol. 26, no. 07, p. 1650025, 2016.

[6] H.-I. Suk, C.-Y. Wee, S.-W. Lee, and D. Shen, “State-space model with deep learning for functional dynamics estimation in resting-state fmri,” *NeuroImage*, vol. 129, pp. 292–307, 2016.

[7] H. Huang, X. Hu, J. Han, J. Lv, N. Liu, L. Guo, and T. Liu, “Latent source mining in fmri data via deep neural network,” in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016, pp. 638–641.

[8] S. Andermatt, S. Pezold, and P. Cattin, “Multi-dimensional gated recurrent units for the segmentation of biomedical 3d-data,” in *Deep Learning and Data Labeling for Medical Applications*. Springer, 2016, pp. 142–151.

[9] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, “Lung pattern classification for interstitial lung diseases using a deep convolutional neural network,” *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[10] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[11] R. Shokri and V. Shmatikov, “Privacy-preserving deep learning,” in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*. ACM, 2015, pp. 1310–1321.

[12] Y. Aono, T. Hayashi, L. Wang, S. Moriai *et al.*, “Privacy-preserving deep learning via additively homomorphic encryption,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1333–1345, 2018.

[13] B. Hitaj, G. Ateniese, and F. Perez-Cruz, “Deep models under the gan: information leakage from collaborative deep learning,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2017, pp. 603–618.

[14] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, “Deep learning with differential privacy,” in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 308–318.

[15] B. Balle and Y.-X. Wang, “Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising,” *arXiv preprint arXiv:1805.06530*, 2018.

[16] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, “Analyze gauss: optimal bounds for privacy-preserving principal component analysis,” in *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*. ACM, 2014, pp. 11–20.

[17] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in *Theory of cryptography conference*. Springer, 2006, pp. 265–284.

[18] F. McSherry and K. Talwar, “Mechanism design via differential privacy,” in *Foundations of Computer Science, 2007. FOCS’07. 48th Annual IEEE Symposium on*. IEEE, 2007, pp. 94–103.

[19] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, “Our data, ourselves: Privacy via distributed noise generation,” in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2006, pp. 486–503.

[20] C. Dwork, A. Roth *et al.*, “The algorithmic foundations of differential privacy,” *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[21] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, “Identifying medical diagnoses and treatable diseases by image-based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[22] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.