Big Data Privacy Protection Technology Integrating CNN and Differential Privacy

Yanfeng Liu*, Ping Li, Min Zhang, Qinggang Liu
School of Information Engineering, Shaanxi Xueqian Normal University, XI'an 710100, China

***Abstract*—To solve the difficulty of balancing privacy and availability in big data privacy protection technology, this study integrates the powerful feature extraction ability of convolutional neural network models with the efficiency of differential privacy technology in data privacy protection. An innovative privacy protection method combining gradient adaptive noise and adaptive step size control is proposed. The experiment findings denote that the research method outperforms existing advanced privacy protection technologies in terms of performance, with an average accuracy of 97.68% and a performance improvement of about 20% to 30%. In addition, for larger privacy budgets, increasing the threshold appropriately can further optimize the effectiveness of research methods. This indicates that through refined noise control and step size adjustment, not only can the privacy protection process be optimized, but also the high efficiency and accuracy of data processing can be maintained. In summary, while ensuring data utility, research methods can not only significantly reduce the risk of privacy breaches, but also optimize privacy protection mechanisms, achieving an ideal balance between protecting personal privacy and maximizing data utility. This innovative approach provides an efficient probability distribution function solution for the field of privacy protection, with the potential to promote further development of related technologies and applications.**

***Keywords*—Convolutional neural network; differential privacy; adaptive noise addition; big data; privacy protection**

I. INTRODUCTION

With the advent of the big data era, data privacy protection has become an increasingly prominent issue. Domestic and foreign researchers have also conducted multiple studies on privacy protection from an academic perspective. Among them, the Convolutional Neural Network (CNN) model has developed rapidly in recent years and made significant progress in privacy protection fields such as image and speech recognition. However, CNN models often rely on massive data during the training process, which may contain sensitive information and can easily lead attackers with different background knowledge to steal improper benefits by directly accessing raw data or indirectly inferring model parameters [1-2]. To address the risk of data privacy leakage faced by CNN models in practical applications, researchers have adopted various technical means to improve CNN models. For example, Zaimi R et al. proposed a deep learning method for detecting phishing websites using a CNN model to address the network threats posed by phishing attacks. The experiment findings indicated that one-dimensional CNN performed well in phishing detection, with an accuracy rate of up to 96.76% [3]. However, this method mainly targets specific types of attacks and does not address the data privacy leakage problem commonly faced by CNN models during the training. Kou X et al. proposed a privacy protection scheme using edge detection technology and CNN model to address the issue of image data leakage, to find a balance between protecting user privacy and ensuring data availability. The outcomes denoted that using edge detection technology for noise addition and feature processing could effectively prevent the leakage of sensitive information in images without sacrificing their practicality [4]. However, this scheme is only applicable to image data and does not consider the privacy protection needs of the model during the training process. Shi J et al. proposed a homomorphic encryption framework based on effective integer vectors to protect the privacy of users in binary CNN models. The outcomes denoted that the training accuracy of this method on the MNIST dataset reached 93.75% [5]. Although the method performs well on specific datasets, it has a large computational overhead and is difficult to scale to large-scale datasets and complex models.

Differential Privacy (DP) is another privacy protection method different from CNN models. This method mainly ensures that even in the event of a data breach, it is impossible to trace specific personal identity information by introducing randomness into the data or algorithm, thereby protecting personal privacy from being leaked [6]. The core of this method is to inject noise into the dataset, reduce the impact of a single data record on the analysis results, and maintain the security of personal information [7]. At present, DP technology has been widely applied in big data environments, especially in data processing and analysis on cloud platforms [8]. For example, the US Census Bureau adopted DP technology to process data in the 2020 census to ensure that personal privacy will not be disclosed while providing statistical information [9]. However, the traditional DP technique has limitations in privacy budget allocation and noise addition mechanism, which can easily lead to data utility degradation and model performance loss. To reduce the risk of supply chain related data information leakage caused by traditional DP technology, Liu M et al. introduced the relevant DP mechanism of logistic regression model and proposed a new supply chain feature selection scheme. Experiments showed that this scheme not only effectively protected the privacy of supply chain data, but also improved data utilization efficiency and enhances prediction accuracy [10]. However, the method is mainly applicable to structured data, and it is difficult to be directly applied to unstructured data (e.g., images, text, etc.). Ma T et al. proposed a DP mechanism for publishing synthetic trajectory database data to enhance the utility of published trajectory data while protecting privacy. The outcomes denoted that this method outperformed other feature-based trajectory synthesis methods in terms of data utility,

achieving a balance between privacy and utility under strict privacy protection [11]. However, the adaptability and robustness of the method in dynamic data environments still need to be further verified.

In summary, although CNN models and DP techniques have made some progress in various privacy protection domains, there are still the following knowledge gaps: (1) Existing methods are inadequate in balancing privacy protection and data availability, and it is difficult to satisfy the needs of high privacy protection strength and high data utility at the same time; (2) The traditional DP techniques lack flexibility in privacy budget allocation and noise addition mechanism, which can easily lead to model performance degradation; (3) Existing schemes mostly target specific data types or attack scenarios, and lack versatility and robustness; (4) In dynamic data environments and diversified attack scenarios, the adaptability and stability of the existing methods need to be improved urgently. Researchers at home and abroad have adopted various technical means, such as edge detection techniques, homomorphic encryption frameworks, and logistic regression models to optimize the CNN model and DP technology to enhance the privacy protection capability of the CNN model and DP technology. These approaches still cannot fully satisfy the needs of different users for balancing privacy and usability in the field of big data privacy protection. To address the above problems, the study intends to fill the knowledge gaps in the following aspects: firstly, a gradient adaptive noise addition model is proposed based on CNN-DP, which solves the balance between privacy protection and data availability by adaptively allocating the privacy budget and optimizing the noise addition mechanism; secondly, an adaptive step-size privacy protection model is designed based on CNN-DP, which draws on the Polyak step-size updating idea and nonlinear extension of constraints based on passive attack algorithm to solve the convergence problem of the model due to privacy protection measures; finally, the proposed method is experimentally verified for its versatility and robustness under diverse datasets and attack scenarios, providing a new solution for the field of big data privacy protection. This research is divided into three sections. The first section describes how the CNN model was improved and how the optimal design model was built, respectively, the second section is a performance test of the new model, and the last section is a summary of the article.

II. Methods and Materials

A. Construction of Gradient Adaptive Denoising Model Based on CNN-DP

During the training, CNN models mainly focus on extracting information from the overall data distribution and do not particularly pay attention to individual data items [12]. Similarly, DP technology pays more attention to the overall statistical information of data after privacy protection when processing data publishing [13]. This consistency in data processing objectives provides a solid theoretical foundation for the combination of DP technology and CNN models. In addition, the training of CNN models requires high computational and communication resources, while DP, as a lightweight algorithm, the combination of the two can achieve complementary advantages [14]. Therefore, the study integrates DP algorithm with CNN model to achieve privacy protection in big data environment. However, the loss function of CNN models will slowly decrease during the convergence, and the loss function will affect the updating of parameters, so the parameters will change in a nonlinear and non-uniform form [15-16]. Based on this characteristic, the study ensures that the protective properties of DP are not compromised by allocating privacy budget reasonably in each iteration update. At the same time, by using gradient adaptive denoising, the constraint noise size is introduced to alleviate the overfitting phenomenon that may occur during CNN training, further improving the model's generalization ability. The gradient adaptive denoising process is shown in Fig. 1.

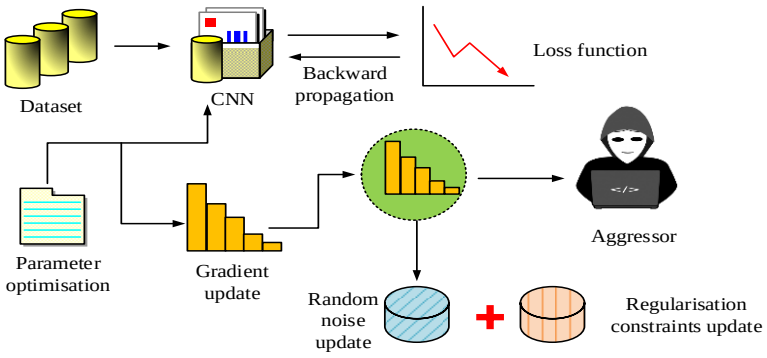

Fig. 1. Gradient adaptive noise injection process.

From Fig. 1, in the gradient adaptive denoising process, the CNN model is first trained routinely, and the input data is processed through forward propagation to calculate the loss function. Subsequently, in the backpropagation stage, the gradient of the loss function with respect to the model parameters is calculated, which reflects the degree of influence of the model parameters on the loss function. At the same time, to introduce DP protection, the study also used Laplace function to add noise to the gradient based on the budget of DP and the sensitivity of the gradient. This addition of random noise helps to protect sensitive information in the training data and prevent attackers from inferring personal information by analyzing the gradient. The gradient after adding noise is used to update the model parameters, and the parameter update rule becomes the original gradient minus the proportionally reduced noise term, where the learning rate determines the size of the step size. Through this approach, the model gradually optimizes in each iteration while ensuring privacy protection. Throughout the

process, gradient adaptive denoising ensures the continuity of model training, while L2 regularization constraints are used to prevent overfitting, enhancing the model's generalization ability and achieving effective model training while protecting privacy. The expression for calculating the L2 regularization term is shown in Eq. (1).

$$L2 = \frac{\lambda}{2n}\sum_{w}\varpi^2 \tag{1}$$

In Eq. (1), λ and n represent the regularization coefficient and sample size, respectively, while ϖ represents the weight parameter. The equation for calculating the loss function C is denoted in Eq. (2).

$$C = C_o + \frac{\lambda}{2n}\sum_{w}\varpi^2 \tag{2}$$

In Eq. (2), C_o represents the original loss function. The expression for gradient update calculation is shown in Eq. (3).

$$\varpi = \varpi - \eta(\frac{\partial C_o}{\partial \varpi} + Lap(\frac{\Delta f}{\varepsilon})) \tag{3}$$

In Eq. (3), η and Δf represent learning rate and global sensitivity, respectively, while ε represents the total privacy budget. The DP privacy protection process is shown in Fig. 2.

In Fig. 2, the core of the DP protection mechanism lies in injecting an appropriate amount of randomness into the data processing process to achieve it. Specifically, for any two adjacent datasets that differ only on one record, applying a random algorithm will result in highly similar probability distributions in their output. Even if individual records are added or deleted from the dataset, the changes in the output results are minimal, effectively reducing the risk of attackers inferring specific individual information based on algorithm outputs. This method provides strong protection for privacy information on the dataset by adding noise value constraints in data queries. The training of the CNN model is indicated in Fig. 3.

In Fig. 3, the training of the CNN model is an iterative process. Firstly, the weights in the network are randomly initialized. In each iteration, the input samples will be passed layer by layer to the network, and the neurons in each layer will multiply the received data with the weights and sum them up. Subsequently, these weighted sums are nonlinearly transformed through activation functions to generate new feature representations. This process is repeated between layers of the network until the network outputs the predicted results. Secondly, the output outcomes are compared with the true labels of the samples and the loss function is calculated. The error signal is then backpropagated back to the network, from the output layer to the input layer, for adjusting the weights of each layer to reduce future errors. By continuously repeating this process, the network weights gradually adjust until the effectiveness of the model on the training data stabilizes, that is, convergence is achieved. The entire process is a manifestation of the stochastic gradient descent algorithm, which relies on the setting of initial weights and updates them in each iteration to optimize the loss function. Due to the correlation between the privacy protection level and privacy budget of DP, this study aims to protect user privacy while ensuring the usability of CNN models as much as possible by adjusting the privacy budget size reasonably.

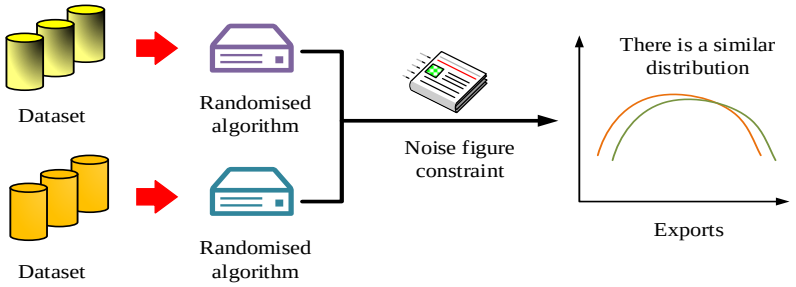

Fig. 2. DP privacy protection process.

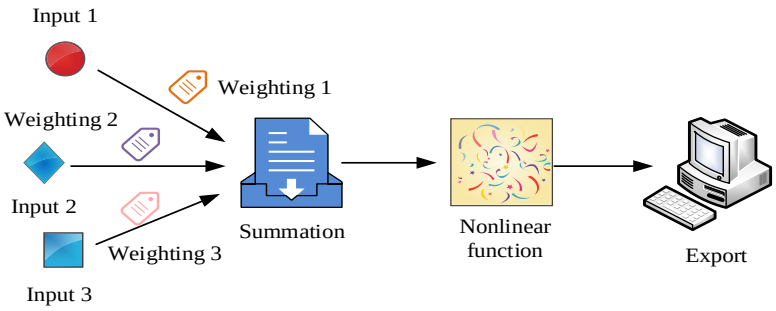

Fig. 3. The training process of the CNN model.

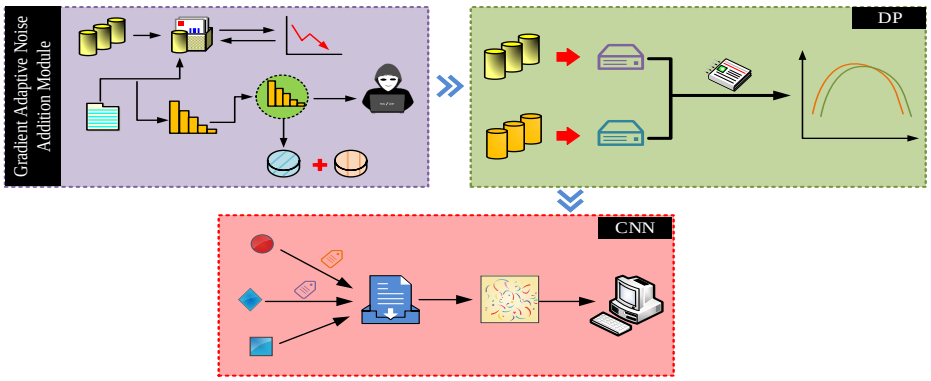

Fig. 4. Overall framework structure of CNN-DP-GAN model.

The privacy budget ε_t calculation equation for the t th iteration is shown in Eq. (4).

$$\varepsilon_t = \varepsilon_1 + (t-1)d \qquad 1 \le t \le T \tag{4}$$

In Eq. (4), ε_1 and d represent the initial privacy budget and the fixed amount of privacy budget added in each iteration, respectively, while T represents the total number of iterations. The equation for calculating the total privacy budget ε after all iterations is denoted in Eq. (5).

$$\varepsilon = T\varepsilon_1 + \frac{T(T-1)d}{2} \tag{5}$$

A CNN-DP Gradient Adaptive Noise (CNN-DP-GAN) model based on CNN-DP was proposed by studying various settings mentioned above. The overall framework structure of the model is denoted in Fig. 4.

In Fig. 4, the CNN-DP-GAN model proposed by the research mainly consists of a gradient adaptive denoising module, a DP privacy protection module, and a CNN training module. The design of this model takes into account the stochastic fine-tuning characteristics of CNN gradient during the training process, and realizes the dynamic allocation of privacy budget during the disturbance process. To prevent excessive noise interference caused by improper privacy budget settings, the model also introduces L2 regularization constraints to regulate the noise level, ensuring a balance between privacy protection and model performance.

B. Construction of an Adaptive Step Size Privacy Protection Model Based on CNN-DP-GAN

Although the CNN-DP-GAN model optimizes the perturbation process by dynamically allocating privacy budgets, effectively balancing privacy protection and data availability, the introduced noise randomness can affect the convergence performance of the model, causing parameters to oscillate when approaching the optimal solution. In addition, the setting of step size parameters is usually complex and susceptible to various factors, resulting in theoretical convergence speeds often being lower than those in practical applications [17]. Therefore, to achieve fast and stable convergence of the model, it is necessary to balance the requirements of privacy protection and the efficiency of model training. To address the convergence issues caused by privacy breaches and noise interference, the CNN-DP-GAN model was nonlinearly extended based on Polyak's step size concept and passive attack algorithm. Relaxation terms were introduced, and stable step size parameters were obtained by combining loss and gradient. By utilizing these measures, a novel adaptive step size privacy protection model based on CNN-DP-GAN was ultimately proposed, namely the CNN-DP-GAN Polyak model.

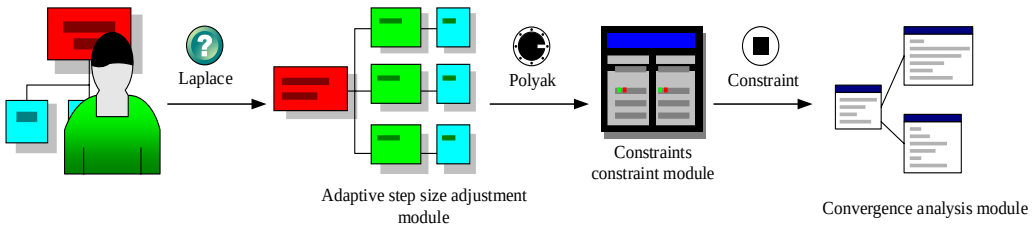

Fig. 5. Overall framework structure of CNN-DP-GAN-Polyak model.

The overall framework structure of the CNN-DP-GAN Polyak model is shown in Fig. 5.

In Fig. 5, the CNN-DP-GAN Polyak model proposed by the research mainly consists of four modules, namely DP privacy protection module, adaptive step size adjustment module, relaxation term constraint module, and convergence analysis module. Among them, the DP privacy protection module is responsible for introducing an appropriate amount of randomness during the model training process, by adding Laplace noise to the gradient or loss function to protect sensitive information in the training data. The adaptive step size adjustment module dynamically adjusts the step size parameters through the Polyak method, redefining the classification update rules for modifying weight vectors at the end of each round to adapt to real-time changes during model training. By monitoring the changes in gradient and loss

function, adaptive step size can more flexibly respond to the convergence behavior of the model, optimize the parameter update process, and improve training efficiency. At the same time, to enhance the robustness and flexibility of the model, the study also introduced relaxation terms to balance the constraints in the optimization process. This constraint helps alleviate overfitting issues and allows the model to maintain sensitivity to data features while meeting privacy protection requirements. The convergence analysis module can ensure that the model can effectively converge to the optimal solution during the iteration process. By analyzing the gradient and parameter update dynamics of the model, the convergence analysis module provides insights into the stability of model training, which helps to understand and predict the behavior of the model and make corresponding adjustments.

However, for most nonlinear models, such as CNN models, the loss function obtained from the output results is often non convex, which makes direct application of the above methods may not be suitable [18]. Therefore, the study also adopted a linearization strategy to handle the loss function, to raise the applicability and optimization efficiency of the model. The equation for calculating the adaptive step size α after linearization is shown in Eq. (6).

$$\alpha = \frac{l_i(w_i)}{\left\|\nabla l_i(w_i) + Lap(\Delta f / \varepsilon_t)\right\|^2} \tag{6}$$

In Eq. (6), $l_i(w_i)$ represents the loss function value at parameter w_i, and $\nabla l_i(w_i)$ represents the gradient of loss function l_i with respect to parameter w_i. The calculation method for the loss function $l(w)$ for classification update is shown in Eq. (7).

$$l(w) = \frac{1}{2m}(\sum_{i=1}^{m}(y^i - h_w(x^i)))^2 \tag{7}$$

In Eq. (7), y^i and $h_w(x^i)$ represent the true labels of the i th sample and the predicted output of the model, respectively, while m represents the number of samples. The calculation expression for the stochastic gradient descent process is shown in Eq. (8).

$$w_j = w_j - \alpha \frac{\partial}{\partial w_i} l(w) \tag{8}$$

In Eq. (8), w_j represents the weight vector. The calculation equation for DP protection of gradient parameters is shown in Eq. (9).

$$w_{t+1} = w_t - \alpha(\nabla l(w_t) + Lap(\frac{\Delta f}{\varepsilon_t})) \tag{9}$$

In Eq. (9), w_{t+1} and w_t represent the model parameters after the ($t+1$)th and t th iterations, respectively. The calculation expression for the parameter update process is shown in Eq. (10).

$$w^{t+1} = w^t - \frac{l_i(w^t)}{\left\|\nabla l_i(w^t) + Lap(\Delta f / \varepsilon_t)\right\|^2}(\nabla l_i(w^t) + Lap(\Delta f / \varepsilon_t)) \tag{10}$$

The expression for calculating the relaxation term constraint is shown in Eq. (11).

$$s_{t+1} = \max\left\{l_i(w_i) - \lambda\left\|\nabla l_i(w_i) + Lap(\Delta f / \varepsilon_t\right\|^2, 0\right\} \tag{11}$$

In Eq. (11), s_{t+1} represents a non-negative relaxation variable. The neural network architecture and parameters used in the training process of the CNN-DP-GAN-Polyak model are shown in Fig. 6.

In Fig. 6, the study used the classic deep learning framework to train the CNN-DP-GAN-Polyak model, ensuring the efficiency of the training process and the wide applicability of the model. At the same time, accuracy is utilized as a key indicator to assess the effectiveness of the model. By testing the model using a dataset within this framework, the relationship between model accuracy and privacy budget is analyzed.

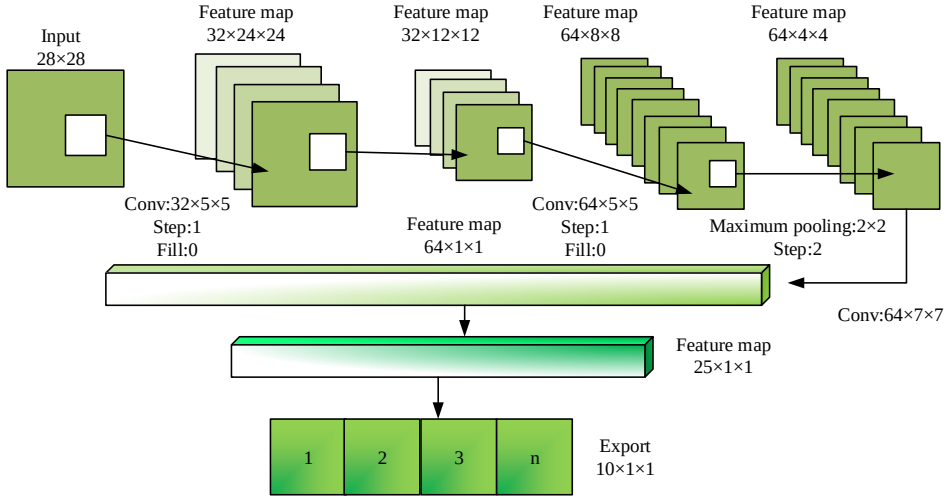

Fig. 6. Neural network architecture and parameters.

III. RESULTS

A. Performance Testing of Gradient Adaptive Denoising Model Based on CNN-DP

To validate the effectiveness of the proposed model, a suitable experimental environment was established. Windows 10 operating system was adopted, equipped with Intel Core i7 CPU, NVIDIA GeForce GPU, 64GB memory, and Python 3.7 programming. The publicly available datasets MNIST, Fashion-MNIST, and CIFAR-10 were utilized as test data sources. These datasets were divided into training and testing sets in an 8:2 ratio. Among them, the MNIST dataset was collected by the National Institute of Standards and Technology in the United States, containing approximately 70000 handwritten grayscale images with a size of 28 × 28. The Fashion-MNIST dataset was provided by a German fashion company and contains 70000 grayscale images of clothing products across 10 categories. CIFAR-10 was a color image dataset containing 10 categories of objects, with an image size of 32 × 32 and a total of 60000 images. These datasets are commonly used benchmark datasets in the fields of machine learning and computer vision, widely used for training and evaluating the effectiveness of models. In addition, parameter selection and optimization are key aspects to ensure model performance. Privacy budget is a core parameter in the DP technique to control the intensity of noise addition, where a smaller privacy budget implies stronger privacy protection but may lead to a decrease in data utility, and a larger privacy budget allows for higher data utility but less privacy protection intensity. The study employed a dynamic privacy budget allocation strategy, where the privacy budget for each iteration was calculated by Eq. (4) and Eq. (5). The noise scale determines the size of the noise added to the gradient, which directly affects the privacy-preserving strength and training stability of the model. The study set the initial and minimum values of the noise scale, and dynamically adjusted the noise size through the gradient adaptive noise addition mechanism. The initial and minimum values of the noise scale were mainly determined through experiments to ensure privacy protection while avoiding excessive noise interference with model training. The specific experimental parameter settings are denoted in Table I.

TABLE I. EXPERIMENTAL PARAMETER SETTING

Serial number	Parameters	MNIST	Fashion-MNIST	CIFAR-10
1	Sample size of batch data	250	256	1500
2	Number of model training rounds	100	100	100
3	Noise scale initial value	2	2	15
4	Noise scale minimum	0.18	0.16	0.10
5	Privacy budget	1	1	1
6	Learning rate	0.001	0.001	0.001
7	Regular term coefficient	0.5	0.5	0.5
8	Gradient trimming value	0.002	0.002	0.002

Based on the parameter settings in Table I, the study first conducted ablation tests on the gradient adaptive denoising model proposed by the research under noisy conditions, with prediction accuracy as the testing indicator. The test results are shown in Fig. 7.

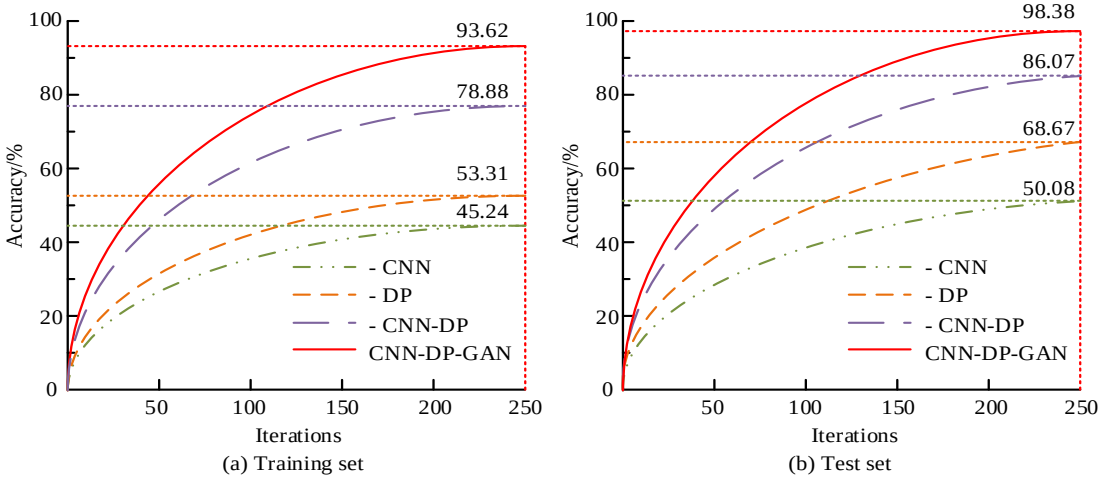

Fig. 7. The ablation test results of the CNN-DP-GAN model.

Fig. 7(a) and Fig. 7(b) show the test results of five modules in the training set and testing set. In Fig. 7(a) and Fig. 7(b), with the increase of iteration times, the prediction accuracy of the five modules showed a steady improvement trend. Among them, the performance of the CNN module was the worst, with a maximum accuracy of only 50.08%. However, when further integrating the DP module and GAN module, the performance of the model was significantly improved. The highest accuracy of the CNN-DP-GAN model reached 98.38%. The reason behind this is that the gradient adaptive denoising method can encourage the model to tend towards selecting better solutions. In this way, the model not only maintained efficient predictive ability while protecting privacy, but also reduced the risk of overfitting through regularization, thereby improving the model's generalization ability. From this, each module component proposed in the study had a positive impact on the final model, which could effectively raise the prediction accuracy of the model. The addition of reasonable noise had

little impact on the accuracy of the CNN-DP-GAN model, and the CNN-DP-GAN model could achieve a balance between privacy and utility on the basis of quantification. In addition, to verify the performance differences between the proposed model and popular models of the same type, the study also introduced the Gradient Descent with Momentum algorithm based on Differential Privacy in CNN (DPGDM), the Differential Private Stochastic Gradient Descent (DP-SGD) based on deep learning and DP, and the Centralized Differential Privacy (CDP) model. The accuracy loss rate of the model was used as the test indicator for comparative testing. The test findings are denoted in Fig. 8.

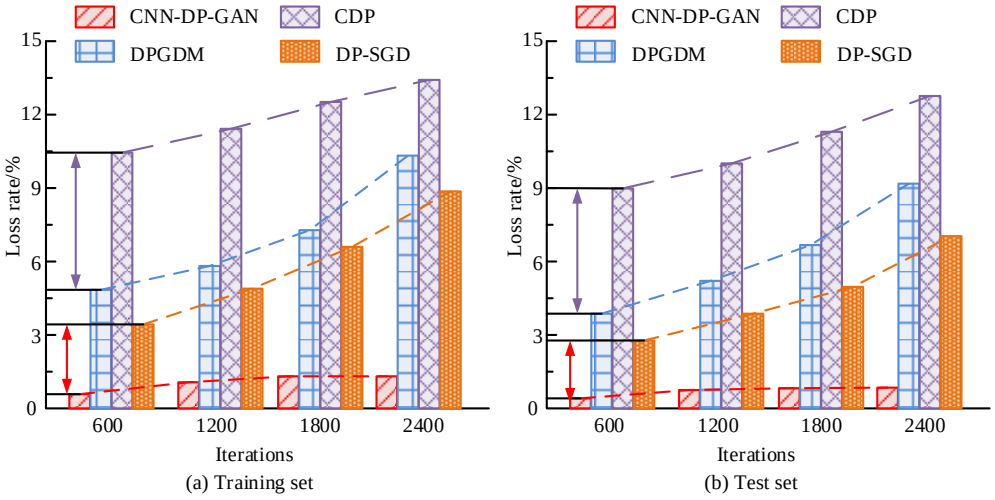

Fig. 8. Accuracy loss rate test results for different models.

Fig. 8(a) showcases the test findings of different models in the training set, and Fig. 8(b) showcases the test findings of different models in the test set. In Fig. 8(a), compared with other models, the CNN-DP-GAN model proposed by the research performed the best. At 600 iterations, the accuracy loss rates of DPGDM, DP-SGD, CDP, and CNN-DP-GAN models were 4.71%, 3.26%, 10.49%, and 1.31%, respectively. This indicated that the CNN-DP-GAN model had significant advantages in maintaining high accuracy and could effectively reduce loss rates. According to Fig. 8(b), at the same number of iterations, the accuracy loss rates of DPGDM, DP-SGD, CDP, and CNN-DP-GAN models were 4.18%, 2.96%, 9.01%, and 1.08%, respectively. These results confirmed that the gradient adaptive denoising method not only had advantages in maintaining model performance, but also continuously optimized the loss rate during the iteration process, further enhancing the privacy protection ability of the model without sacrificing accuracy excessively. This strategy provides an effective technical means for achieving efficient and accurate data processing while protecting privacy.

B. *Performance Testing of Adaptive Step Size Privacy Protection Model Based on CNN-DP-GAN*

When using Laplace mechanism for privacy protection, privacy budget and sensitivity are key factors affecting the level of privacy protection. Therefore, the research mainly focused on these two core variables and explored how to achieve the optimal balance between model privacy protection and utility. The sensitivity and privacy budget values under different iteration times are shown in Table II.

Due to the Laplace perturbation, the variance is equal to the ratio of sensitivity to privacy budget. Therefore, the study controlled the overall privacy budget to remain unchanged. According to Table II, experiments were conducted at different sensitivities to compare the average final accuracy of different models. The test results are indicated in Fig. 9.

TABLE II. SENSITIVITY AND PRIVACY BUDGET TAKES FOR DIFFERENT NUMBER OF ITERATIONS

Datasets	Parameters		Sensitivity	Total budget
MNIST	Different number of iterations	300	0.5	72.5
		600	0.5	141
		1200	0.5	279.5
Fashion-MNIST	Different number of iterations	300	0.5	143
		600	0.5	283.5
		1200	0.5	960
CIFAR-10	Different number of iterations	300	0.5	217.5
		600	0.5	312.5
		1200	0.5	687.6

Fig. 9(a) and Fig. 9(b) show the comparison curve of the average final accuracy of the models in the MNIST, and CIFAR-10 dataset, respectively. In Fig. 9(a), compared with other models, the proposed model achieved better model performance while ensuring a balance between privacy and utility. The average final accuracies of DPGDM, DP-SGD, CDP, and CNN-DP-GAN Polyak models were 70.23%, 82.36%, 86.08%, and 97.68%, respectively. In Fig. 9(b), the CNN-DP-GAN Polyak model proposed by the research performed the best, with an average final accuracy of 92.08%, which was a performance improvement of 20% to 30% compared to other models. From this, it can be seen that under the constraint of data utility, the model could effectively minimize the risk of privacy leakage and optimize the privacy protection mechanism, thereby obtaining a probability distribution function that achieves the best balance between protecting privacy and maintaining data utility. The effectiveness of the research method was proved. Finally, the study also explored the impact of different privacy budgets on the adaptive step size adjustment process. The test results are indicated in Fig. 10.

Fig. 10(a), (b), and (c) show the accuracy variation curves with threshold settings of 0.01, 0.1, and 1 at 300 iterations. From Fig. 10, in the early stages of iteration, when the privacy budget was set to 5, the adaptive step size adjustment method has not fully utilized its advantages, resulting in poor performance of the CNN-DP-GAN-Polyak model. As the iteration progressed, a smaller threshold setting could help improve the performance of the CNN-DP-GAN-Polyak model when the privacy budget was low. On the contrary, for larger privacy budgets, increasing the threshold appropriately could optimize the performance of the CNN-DP-GAN-Polyak model. This indicated that the setting of privacy budget and threshold needed to be dynamically adjusted based on iteration progress and privacy protection requirements to achieve the optimal balance between privacy protection and data utility. Through this meticulous adjustment, it was possible to maximize the predictive accuracy and practicality of the model while minimizing the risk of privacy breaches.

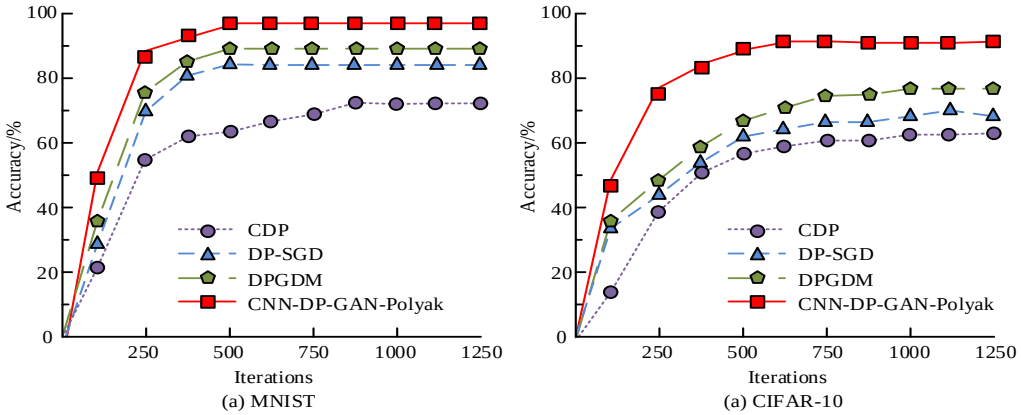

Fig. 9. Comparison curves of final accuracy averages of different models.

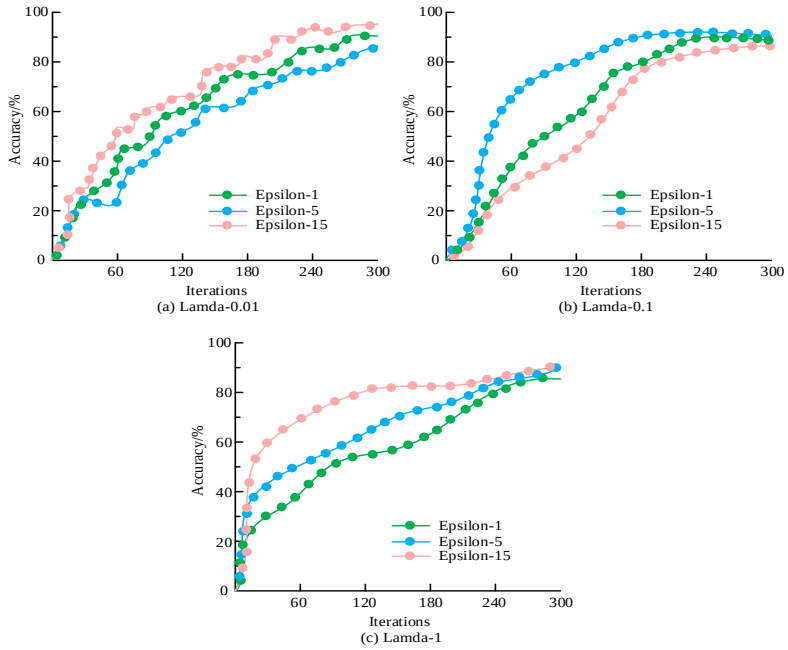

Fig. 10. Accuracy variation curves for different threshold settings.

IV. Conclusion

The rapid development of computer vision largely relies on the innovative construction of deep learning models and the participation of large-scale datasets. With the continuous advancement of technology, data privacy protection has gradually become a hot research topic. In practical application scenarios, CNN models face significant risks of data privacy breaches when handling tasks involving sensitive information. To effectively address this challenge, a novel big data privacy protection technique is proposed by combining CNN models with DP technology, utilizing gradient adaptive denoising method and adaptive step size privacy protection method. The outcomes denoted that the gradient adaptive denoising method could effectively guide the model to choose a better solution. In a noisy environment, the highest accuracy of the CNN-DP-GAN model reached 98.38%, with an accuracy loss rate of only 1.08%. In addition, compared with other advanced models, the CNN-DP-GAN-Polyak model proposed by the research performed the best, with an average final accuracy of 97.68%. As the iterative process progressed, especially with low privacy budgets, appropriate threshold settings have been shown to help improve the performance of the CNN-DP-GA-Polyak model. From this, the method proposed by the research can achieve good model performance while ensuring a balance between privacy protection and data utility. However, research mainly evaluates the performance of models in terms of privacy protection and data utility based on privacy budget and model accuracy. Future work can expand the focus to assess the ability of models to resist attackers with auxiliary background knowledge, thereby comprehensively improving the breadth and depth of model validation.

V. Funding

The research is supported by: Shaanxi Fundamental Science Research Project for Mathematics and Physics (Grant No.23JSY051), Research on Privacy Protection Algorithms in Big Data Computing.

References

[1] Qiang W, Liu R, Jin H. Defending CNN against privacy leakage in edge computing via binary neural networks. Future Generation Computer Systems, 2021, 125(37):460-470.

[2] Ding Y, Shen W, Hai-sheng L, et al. Blockchain Trusted Privacy Service Computing Model for CNN. Acta Electronica Sinica, 2022, 50(6):1399-1409.

[3] Zaimi R, Hafidi M, Lamia M. A deep learning approach to detect phishing websites using CNN for privacy protection. Intelligent decision technologies: An international journal, 2023, 17(3):713-728.

[4] Kou X, Wang F, Zhu H, et al. Masked image: Visually protected image dataset privacy-preserving scheme for convolutional neural networks. Peer-to-Peer Networking and Applications, 2024, 17(4):2523-2537.

[5] Shi J, Zhao X. Anti-leakage method of network sensitive information data based on homomorphic encryption. Journal of Intelligent Systems, 2023, 32(1):2517-39.

[6] Acharya M, Mohbey K. Differential Privacy-Based Social Network Detection Over Spatio-Temporal Proximity for Secure POI Recommendation. SN Computer Science, 2023, 4(7):1-10.

[7] Yuan J , Wang Z , Liu D H .Retracted: Multi-vehicle group-aware data protection model based on differential privacy for autonomous sensor networks. IET circuits, devices & systems, 2023, 17(4):278-290.

[8] Chen Q, Ni Z, Zhu X, et al. Differential privacy histogram publishing method based on dynamic sliding window. Frontiers of Computer Science, 2022, 17(51):1-12.

[9] Waller L A. Global and local impacts of differential privacy on estimates of health care inequity. Health services research, 2022, 57(2):204-206.

[10] Liu M, Song X, Li L W. Correlated differential privacy based logistic regression for supplier data protection. Computers & Security, 2024, 136(12):103542.1-103559.

[11] Ma T, Deng Q, Al-Nabhan R N. A privacy-preserving trajectory data synthesis framework based on differential privacy. Journal of information security and applications, 2023, 77(9):103550.1-103550.11.

[12] Divya, Anand N, Sharma G. Convolutional neural network (CNN) and federated learning-based privacy preserving approach for skin disease classification. The Journal of Supercomputing, 2024, 80(16):24559-24577.

[13] Rania Z, Mohamed H, Mahnane L. A deep learning approach to detect phishing websites using CNN for privacy protection. Intelligent Decision Technologies, 2023,17(3):713-728.

[14] Fan Z, Zhi L, Hao W. PPCNN: An efficient privacy-preserving CNN training and inference framework. International Journal of Intelligent Systems, 2022, 37(12):10988-11018.

[15] Weizhong Q, Renwan L, Hai J. Defending CNN against privacy leakage in edge computing via binary neural networks. Future Generation Computer Systems, 2021, 125(12):460-470.

[16] Zhang J, Si K, Zeng Z, et al. IEA-DP: Information Entropy-driven Adaptive Differential Privacy Protection Scheme for social networks. The Journal of Supercomputing, 2024, 80(14):20546-20582.

[17] Gopahanal Manjunath M, Vyjayanthi C, Modi C N. Adaptive step size based drift-free P&O algorithm with power optimiser and load protection for maximum power extraction from PV panels in stand-alone applications. IET renewable power generation, 2021, 15(6):1270-1285.

[18] Choudhuri S, Adeniye S, Sen A. Distribution Alignment Using Complement Entropy Objective and Adaptive Consensus-Based Label Refinement For Partial Domain Adaptation[C]//Artificial Intelligence and Applications. 2023, 1(1): 43-51.