Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Noise Addition Strategies for Differential Privacy in Stochastic Gradient Descent

Kangjie Lu

China-Japan Software Institute, Dalian University of Technology, Dalian, China

2029160964@mail.dlut.edu.cn

Abstract. Differential privacy technology is more and more widely used in the field of machine learning, especially in the gradient descent algorithm (SGD). Protecting data privacy by adding noise has become a hot topic of research. This paper reviews the noise addition strategy of differential privacy SGD from multiple dimensions, including adjustment based on noise distribution, adjustment based on gradient norm, adjustment based on privacy budget, and method based on model architecture. Each strategy has different performances in terms of privacy protection level, model performance loss and computational complexity. This article compares and analyzes these differences in detail, aiming to provide valuable reference for researchers and practitioners. This article also discusses how to combine federal learning and differential privacy technology to protect data privacy more efficiently in a secure multi-party computing (MPC) environment. Through the review of this article, we can see the wide application of differential privacy in machine learning and deep learning and its importance in the field of privacy protection. At the same time, we also show the direction and challenges of future research.

Keywords: noise addition strategies; differential privacy; stochastic gradient descent.

1. Introduction

In today's data-driven world, the training and deployment of machine learning models involves the processing of a large amount of sensitive data, which has aroused widespread attention to data privacy protection. As a powerful privacy protection mechanism, Differential Privacy (DP) ensures that the privacy of individual data is not leaked by introducing noise into the data or model training process. Especially in the core training method of deep learning, Stochastic Gradient Descent (SGD), how to effectively integrate differential privacy has become an important research topic. As an effective privacy protection mechanism, differential privacy is widely used in the training process of machine learning and deep learning models. AtIn differential privacy SGD, how to balance the relationship between privacy protection and model performance is a hot topic and challenge in current research.

This article summarizes the noise addition strategy of differential privacy SGD and compares the differences between different strategies in terms of privacy protection level, model performance loss and computational complexity. Through in-depth analysis of advantages and disadvantages, it provides reference for researchers and practitioners. At the same time, it discusses how to effectively apply differential privacy to protect data privacy in scenarios such as federal learning and MPC. The specific logical block diagram is shown in Figure 1.

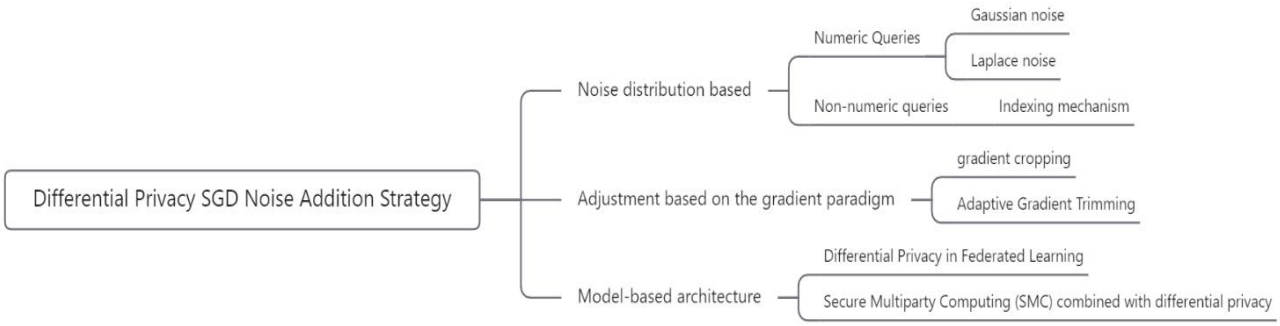

Figure 1. Differential Privacy SGD Noise Addition Policy (Photo/Picture credit :Original)

2. Based on noise distribution

Among these strategies, Gaussian noise addition and Laplace noise addition are the two most common methods based on noise distribution. They are particularly widely used in numerical queries, but when dealing with non-numerical queries, the exponential mechanism shows its unique advantages. While providing privacy protection, each method will also have a different degree of impact on model training and final performance.

2.1. Numerical query

1) Gaussian noise: Gaussian noise and its variants are widely used in SGD and have a great impact on differential privacy training performance, especially in privacy protection and reduced overfitting scenarios.

First of all, according to the research of literature [1], even in the case of a powerful Gaussian disturbance (Noisy-SGD), the implied deviation in SGD still exists. This study reveals that the L2 norm of the added Gaussian noise continues to exceed the gradient itself, which significantly affects the training trajectory. This discovery points out that in SGD, a specific gradient noise geometry shows a certain degree of robustness to the Gaussian disturbance, and thisThe difference in shape may lead to different implied deviations.

Further, the literature [2] discusses the application of the general Gaussian noise mechanism in protecting data privacy. This method adds Gaussian noise to the data.It aims to protect data effectively through the introduction and application of trp(M) while protecting privacy or reducing overfitting.. This illustrates the practical application and importance of Gaussian noise in the field of data privacy.

Although Gaussian noise and its variants show great potential in privacy protection and reduction of overfitting, they also bring new challenges to the model training process, especially in model optimization and performance. These challenges need to be overcome through further research and innovative approaches.

2) Laplace noise: Laplace noise is a technology commonly used for privacy protection, which is particularly important when processing data containing sensitive information. It is based on the Laplace distribution, which is a zero-centered bilateral exponential distribution, and its shape is determined by position parameters and scale parameters. The introduction of Laplace noise can effectively confuse sensitive information in the data, thus providing privacy protection without significantly affecting the overall statistical characteristics of the data.

First of all, the literature [3] shows the application of the concept of local differential privacy (LDP) by discussing the estimation of drift coefficient parameters in the random differential equations (SDEs) system. Under the constraint of only using and publishing public data, by adding an appropriate amount of Laplace noise, it not only protects data privacy, but also accurately estimates key parameters. This process makes full use of the effectiveness of Laplace noise in privacy protection. It provides a useful exploration for the application of privacy protection technology in complex system analysis.

Further, the literature [4] emphasizes the importance of the Laplace mechanism in the framework of differential privacy in response to the privacy protection challenges in the Internet of Things environment. By adding Laplace noise to sensitive data, this mechanism shows the optimal effect in protecting privacy based on the $\ell 1$ norm. At the same time, considering the protection needs of high-dimensional data, a variant of the Laplace mechanism is proposed to optimize the mean square error. This indicates the noise of Laplace.It can not only adapt to different privacy protection needs, but also achieve optimal protection in diversified application scenarios.Through these examples, we can see the wide application of Laplace noise in the field of privacy protection and its gradual and in-depth research. From simple privacy protection to complex system analysis to dynamic data processing, the flexibility and effectiveness of Laplace noise have been fully demonstrated and verified, reflecting its important value in protecting personal privacy and data security.

2.2. Non-numeric query

Index mechanism is an algorithm widely used in the field of differential privacy, which aims to provide accurate query results while protecting data privacy. It works by assigning a probability to each possible output, which is proportional to the utility of the output. In differential privacy, the utility function is usually used to measure the "closeness" of the output to the actual data set, while the exponential mechanism optimizes the accuracy of the query results by selecting the output of efficient values. The key of this mechanism is to adjust these probability distributions to ensure that even if noise is added to the output, the usefulness of data can be retained to the greatest extent while protecting individual privacy.

In addition, the discussion of the exponential mechanism in the literature [5] further deepens our understanding of its efficiency in dealing with large-scale output space. It is pointed out that this mechanism randomly selects the output based on the probability distribution defined by the scoring function, which not only protects data privacy, but also can be selected according to the quality of the data. In the face of large-scale output space, efficient sampling can be achieved by adopting a special structure or method (such as the Markov chain Monte Carlo (MCMC) method), which is particularly important for practical application. Through the gradual deepening of these examples, we can see that the application of the index mechanism in the field of differential privacy is diversified and flexible, which can not only effectively protect personal privacy, but also ensure the effective use of data. From basic principles to specific application cases, the index mechanism shows its great potential to achieve a balance between protecting privacy and maintaining data quality.

3. Adjustment based on gradient norm

As an adjustment method based on the gradient norm, gradient clipping reduces the amount of noise that must be added by limiting the maximum value of the gradient, which is a key technology in differential privacy SGD. The adaptive gradient cutting strategy further optimizes this process and dynamically adjusts the cutting threshold according to the performance of the model in the training process, aiming to better balance the performance of the model and privacy protection.

3.1. Gradient clipping

Gradient cutting is an important technology in the optimization algorithm, which aims to prevent the gradient explosion by limiting the size of the gradient, which is conducive to the stability and effective training of the model. In the learning environment of privacy protection, the role of gradient cutting is not only to control the scale of the gradient, but also to reduce the reduction of training efficiency that may be introduced due to the addition of noise. This article will combine several literature examples to discuss the application of gradient cutting in different scenarios and its impact on training performance.

First of all, the literature [1] mentioned that even in the scenario of no gradient cropping (only noise addition), the performance degradation problem of large-scale training still exists. This observation shows that the performance decline is not only caused by gradient cutting, but also the result of the added noise and the randomness of SGD itself. This suggests that when designing privacy protection algorithms, we need to comprehensively consider the effects of noise addition and gradient clipping, and seek balance.

Further, in the literature [6], the FedFDP algorithm introduced shows how to meet the requirements of fairness by designing a special gradient cutting strategy. Unlike the traditional gradient cutting method, FedFDP considers the fairness between different clients in the training process, and may need to adopt different standard gradient cutting for different clients. This method not only ensures that the gradient scale is controlled, but also promotes fairness, butIt may increase the complexity and execution difficulty of the algorithm.

3.2. Adaptive gradient cutting

Adaptive gradient clipping is an advanced technology that protects differential privacy in deep learning. It aims to optimize the training process of the model by dynamically adjusting the clipping threshold while minimizing the risk of privacy leakage. This method allows the model to find a more balanced point between improving performance and protecting user privacy.This paper will explore the application of adaptive gradient cutting in different studies and its impact on the model training process.

First of all, the adaptive tailoring strategy proposed in the literature [7] is to calculate the gradient of each parameter in each iteration, and then dynamically determine the cutting threshold according to the actual training of the model. This method allows the model to reduce the risk of leakage of privacy while maintaining gradient effectiveness and training stability. By optimizing the size of the gradient before adding noise, adaptive gradient clipping helps to find a better balance between privacy protection and model performance. In addition, adaptive adjustment of the cutting threshold can reduce information loss and improve training efficiency, which is a significant advantage. However, implementing this strategy may require additional computing resources to monitor and adjust the cropping threshold, which may be a potential disadvantage.

Then, the literature [6] introduces the method of adaptive cropping of the upload loss value in the FedFDP algorithm. This strategy aims to protect data privacy while maximizing the utility of the model by adaptively adjusting the cutting threshold of the loss value. This method balances privacy protection and model performance, which is an improvement. However, its challenge is how to accurately set and adjust the cutting threshold of the loss value in order to achieve the best balance between model utility and privacy protection.

Finally, the FedDPA framework proposed in the literature [8] uses adaptive gradient clipping to deal with personalized parameters and shared parameters. This differentiation processing strategy is to reduce the loss of information to personalized parameters while protecting the privacy of the overall model. Through real-time monitoring of the training and parameter update of the model, this adaptive method not only improves the efficiency of privacy protection, but also optimizes the gradient cutting parameters, thus accelerating the convergence of the model and improving the training efficiency. This shows that in personalized federal learning, by flexibly adjusting the cutting strategy, the problems of lack of flexibility and convergence can be effectively solved, while improving the performance and robustness of the model.

4. Based on model architecture

In the federal learning environment, the application of differential privacy is facing new challenges and opportunities. By updating the model locally and adding noise, and then sending the updated gradient back to the central server for aggregation, Federal Learning provides an efficient method of using decentralized data for model training, and also provides an additional layer of protection for data privacy.

4.1. Differential privacy in federal learning

Federal learning is a machine learning paradigm that allows multiple clients to train models together without sharing their original data. This helps to overcome the problem of data islands while protecting user privacy. However, there are two main problems.. The first is the issue of fairness,Global models trained on different clients may show better or worse performance on some clients than other clients, resulting in unfair results.The second isPrivacy disclosure issues,Even if the data is not shared directly, sensitive information may be leaked through transmission gradients or model updates (for example, member inference attacks). In order to solve the problem of fairness,A federal learning algorithm called FedFair is proposed [6]. The algorithm aims to ensure that the performance of models on different clients is more fair by adjusting the training process. Further, the literature [9] discusses the number of times the parameter server (PS) queries the client and the number of client reply queries, so as to maximize the accuracy of the final model. This paper examines two commonly used differential

privacy mechanisms, the Laplace mechanism and the Gaussian mechanism. Through the analysis of convergence rate, the optimal number of queries and replies in federal learning that introduces differential privacy is determined, so as to maximize the accuracy of the final model. The experimental results show that properly setting the number of queries and replies can significantly improve the accuracy of the final model of federal learning under differential privacy while protecting privacy.

In order to solve the problem of privacy leakage at the same time, the FedFDP algorithm is proposed based on FedFair and the concept of differential privacy is introduced. FedFDPThe flowchart is shown in figure 2 [10]

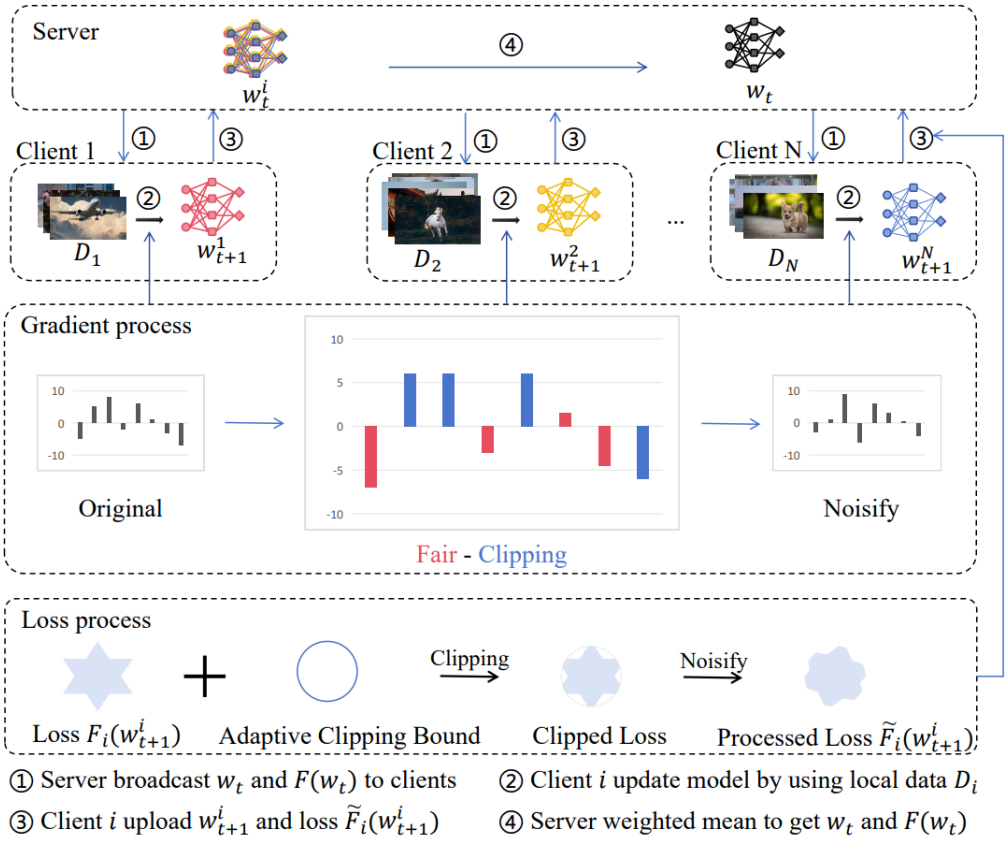

Figure 2. FedFDP framework [10]

Subsequently, it is mentioned in the literature [10]. A novel scheme that combines differential privacy protection technology and federal learning is called Fed-SMP. This scheme aims to solve a major challenge in achieving differential privacy protection in federal learning: the negative impact of DP noise on model accuracy, especially for deep learning models with many parameters. The Fed-SMP scheme mitigates the impact of privacy protection on model accuracy through a new technology, Sparsified Model Perturbation (SMP). In this scheme, the local model is first sparsed, and then disturbed by Gaussian noise. This method not only provides customer-level differential privacy protection, but also saves communication costs by reducing the amount of data to be transmitted.

The existing DPFL method may cause the loss landscape to become steeper and less robust to weight disturbances, resulting in a serious performance degradation. DP-FedSAM, a new differential privacy federal learning algorithm, is mentioned in the literature [11], which aims to defend against inference attacks and mitigate the leakage of sensitive information in federal learning. DP-FedSAM reduces the negative impact of DP by using gradient disturbances. Specifically, the algorithm uses the Sharpness Aware Minimization (SAM) optimizer to generate local smoothness with better stability and weight disturbance robustness. The model improves the performance of the model. Compared with the existing differential privacy federal learning benchmark, the DP-FedSAM algorithm achieves the most advanced (SOTA) performance.

4.2. SMC combined with differential privacy

In the field of modern data analysis and machine learning, federal learning is a new type of distributed learning paradigm that allows multiple participants to train models together without sharing their original data. This method plays a key role in improving data privacy protection. However, in order to further strengthen data privacy protection, differential privacy technology is often used in federal learning. Differential privacy protects personal privacy by adding a certain amount of noise to the data to prevent any single data point from being identified. Although this method effectively improves the level of privacy protection, it may also have a certain impact on the accuracy of the model. Therefore, researchers seek more efficient privacy protection methods, and secure MPC has become one of the important technical means.

Secure MPC is a technology that allows multiple parties (or "participants") to jointly calculate the result of a function (for example, the aggregation of model updates) without exposing their respective inputs (i.e., local model updates). The literature [12] mentions PrivatEyes, a privacy-enhanced training method for appearance-based gaze estimation, which combines federal learning and secure MPC technology. In PrivatEyes, MPC is used to ensure that individual attention data remains private even when most aggregation servers may be malicious. This method allows personal data to be protected even when most aggregation servers may not be trusted, while maintaining the accuracy of estimates without a significant increase in calculation costs.

Further, a study described in the literature [13], in which a new privacy protection machine learning protocol has been developed to train linear regression, logical regression and neural network models by using secure MPC. The study focuses on model training using the stochastic gradient descent method while ensuring that data privacy is protected. The protocol is based on a two-server model, in which data owners distribute their private data to two non-conspiracy servers. These servers use secure two-party computing (2PC) to train various models on joint data without knowing each other's specific data content. The research has also developed new technologies to support safe arithmetic operations on shared decimal numbers. This is essential for accurate numerical calculation while protecting privacy.

Secure MPC provides an effective solution for privacy protection in federal learning, which not only strengthens data privacy protection, but also maintains the accuracy of model training. Through the examples of PrivatEyes and other related studies, we can see the potential of MPC technology in improving the level of privacy protection and the challenges that may be encountered in practical application. In the future, with the further development of computing technology, these challenges are expected to be solved, so that MPC can play a greater role in a wider range of scenarios.

5. Analysis and Comparison of Differential Privacy SGD Noise Addition Strategy

By comprehensively considering the differential privacy SGD noise addition strategy of multiple dimensions, we can find that there are significant differences between different strategies in terms of privacy protection ability, impact on model performance and implementation difficulty. In order to show these differences more intuitively, this article will select key indicators, such as privacy protection level, model performance loss, computational complexity, etc., for in-depth comparison and analysis. The specific analysis is shown in Table I.

Table 1. Differential Privacy SGD Noise Addition Strategy Analysis and Comparison

Method category	Literature	Level of privacy protection	Loss of model performance	Calculated complexity	Advantage	Shortcoming
Gaussian noise addition	DPSUR, DPSGD related research	Medium to high	Medium	Medium	Simple and easy to implement, suitable for a variety of situations	It is necessary to choose the appropriate standard deviation, which has a certain

						impact on the performance of the model.
Laplace noise addition	Research on Differential Privacy Protection of Local Fingerprint Images	Tall	Medium to high	Medium	High requirements for privacy protection	Adding more noise has a great impact on the performance of the model.
Index mechanism	Shifted Interpolation, etc.	Tall	Variable	Variable	It is suitable for non-numerical queries and provides a high level of privacy protection.	The impact on the performance of the model is uncertain, and the computational complexity is high.
Gradient cutting	FedFDP, Clip21, DPFormer	Medium	Low to medium	Low to medium	It can reduce the amount of added noise and reduce the loss of model performance.	The cutting threshold needs to be adjusted, which may lead to information loss.
Adaptive gradient cutting	FedFDP, adaptive differential privacy personalized federal learning	Tall	Low to medium	Medium	Dynamically adjust the cutting threshold to balance privacy protection and model performance.	It is necessary to monitor the model performance in real time, and the computational complexity is high.
Differential privacy in federal learning	FedFDP, DP-FedSAM	Tall	Low to medium	High (due to aggregation and communication expenses)	Protect multi-party data and improve the level of privacy protection	The aggregation and communication cost are large, and the computational complexity is high.
SMC combined with differential privacy	PrivatEyes	Tall	Low	Tall	Multi-party participation in training models to protect data privacy	Secure MPC technical support is required, and the computing complexity is high.

The specific level of privacy protection, performance loss and computational complexity will vary according to different experimental settings, data sets, task types and other factors.SoIt is usually necessary to weigh the applicability and effect of different strategies according to the needs of specific application scenarios. For example, in scenarios with extremely high privacy protection requirements, the Laplace noise addition or index mechanism may be preferred, while in federal learning scenarios, considering the heterogeneity of data distribution and communication costs, it may be more inclined. Use the differential privacy policy in federal learning.

6. Conclusion

This paper comprehensively analyzes the application of differential privacy technology in random SGD from multiple angles, paying special attention to the performance of noise addition strategy in privacy protection, model performance and computational complexity. By comparing Gaussian noise, Laplace noise addition, exponential mechanism, gradient cutting, adaptive gradient cutting, and applying differential privacy in a federal learning and secure MPC environment, this paper reveals the advantages and disadvantages of different strategies and applicable scenarios. In addition, it also discusses how to minimize the impact on model performance while protecting privacy, and the computational complexity that needs to be considered when implementing these methods.

Through in-depth analysis and comparison, this article provides researchers and practitioners with valuable insights on how to choose a differential privacy SGD strategy suitable for the needs of specific scenarios. Especially in today's increasingly important consideration of data privacy protection, the choice and application of differential privacy technology has become particularly critical. Future research may pay more attention to how to optimize these strategies to achieve higher privacy protection and lower model performance losses, while reducing computational complexity. In addition, with the development of technology, exploring new differential privacy policies and combining differential privacy technology with other privacy protection technologies will also be a direction worth studying.

References

[1] T. Sander T, M. Sylvestre, A. Durmus Implicit Bias in Noisy-SGD: With Applications to Differentially Private Training. arXiv preprint arXiv:2402.08344, 2024.

[2] A. Nikolov, H. Tang. General Gaussian Noise Mechanisms and Their Optimality for Unbiased Mean Estimation. arXiv preprint arXiv:2301.13850, 2023.

[3] C. Amorino, A. Gloter, H. Halconruy. Evolving privacy: drift parameter estimation for discretely observed iid diffusion processes under L DP. arXiv preprint arXiv:2401.17829, 2024.

[4] F. Koufogiannis, S. Han, G. Pappas Optimality of the laplace mechanism in differential privacy. arXiv preprint arXiv:1504.00065, 201 5, 10.

[5] S. Roy, A. Tewari. On the Computational Complexity of Private High-dimensional Model Selection via the Exponential Mech anism. arXiv preprint arXiv:2310.07852, 2023.

[6] X. Ling, J. Fu, Z. Chen, et al. FedFDP: Federated Learning with Fairness and Differential Privacy. arXiv preprint arXiv:2402.1602 8, 2024.

[7] Z. Chu, J.He, D. Peng, et al. Differentially Private Denoise Diffusion Probability Model. IEEE Access, 2023.

[8] X. Yang, W. Huang, M. Ye. Dynamic personalized federated learning with adaptive differential privacy. Advances in Neural Information Processing Systems, 2023, 36: 72181-72192.

[9] Y. Zhou, X. Liu, Y. Fu , et al. Optimizing the Numbers of Queries and Replies in Federated Learning with Differential Privacy. ar Xiv preprint arXiv:2107.01895, 2021.

[10] R. Hu, Y. Guo Y, Gong. Federated learning with sparsified model perturbation: Improving accuracy under client-level differe Ntial privacy. IEEE Transactions on Mobile Computing, 2023.

[11] Y. Shi Y, Liu, K. Wei, et al. Make landscape flatter in differentially private federated learning, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 24552-24562.

[12] M. Elfares M, Reisert P, Hu Z, et al. PrivatEyes: Appearance-based Gaze Estimation Using Federated Secure Multi-Party Computation. arXi V preprint arXiv:2402.18970, 2024.

[13] P. Mohassel, Y. Zhang. Secureml: A system for scalable privacy-preserving machine learning, 2017 IEEE symposium on secur Ity and privacy (SP). IEEE, 2017: 19-38.