Memristor-Based Neuromorphic Hardware Improvement for Privacy-Preserving ANN

Jingyan Fu, *Student Member, IEEE*, Zhiheng Liao, *Student Member, IEEE*, and Jinhui Wang, *Senior Member, IEEE*

***Abstract*— Because of collecting a large amount of personal data, when the artificial neural network (ANN) is used in human-related topics, it has raised great concerns on privacy preservation. A robust solution is to introduce a noise injection mechanism as differential privacy that promises strong theoretical privacy guarantees. However, privacy-preserving ANN with noisy input data has a substantial risk of reducing the recognition accuracy. Therefore, it is urgently needed to have technologies that can make users' data applied to neural networks while strictly protecting sensitive information. In this paper, a linear optimization (LO) method is proposed to address this accuracy degradation by optimizing the performance of memristor in weight updating processes. Instead of complying with the traditional hardware and algorithm, the LO method calculates update parameters along a piecewise line by using different input pulses. The proposed method can mitigate the nonlinear problem of memristor without prereading the precise current conductance each time, thereby avoiding complex peripheral circuits. The effectiveness of the proposed LO method with two-segment, three-segment, and four-segment models is investigated, respectively. The results show that under different nonlinearity and different perturbation noise required by differential privacy theory, the LO method can increase the recognition accuracy of Modified National Institute of Standards and Technology (MNIST) handwriting digits by 39.67% on average, which provides more space and margin for privacy-preserving technology.**

***Index Terms*— Artificial neural network (ANN), memristor, neuromorphic hardware, nonlinearity, privacy preservation.**

I. Introduction

ARTIFICIAL neural networks (ANNs) have been applied successfully in a broad range of applications such as computer vision, speech recognition, machine translation, robot control, and medical diagnosis [1]. The performance of ANN is often directly dependent on a large number of parameters to encode the network and summarize representative data sets and then to build models for new data analysis. However, with the fast development of ANN technology and the dramatically increasing amount of user and their linked databases, ANN encounters two problems. The first one is computational bottlenecking. It occurs when the communication and training processes bring high latency and high energy cost that challenge hardware with von Neumann architecture and CMOS technology. The second one is privacy preservation. The privacy needs to be protected rigorously because the data collected from users are often crowdsourced and may contain sensitive personal information [2]. For example, in Internet of things (IoT) applications, devices are typically dispersed across the globe, allowing for instant communication. The collected data in some cases may contain very private information of users. However, users are usually unaware of how much data is actually collected and how the data is used or shared with others. Therefore, there is always a concern from experts about privacy preservation [3].

To address computational bottleneck, many technologies are explored. Memristor-based neuromorphic computing is such a flexible and attractive technology to meet the increasing needs of data processing [4], [5]. A memristor is a device with only three layers structure that can not only realize desirable device properties such as sub-10-nm feature sizes [6], subnanosecond switching speed [7], [8], long write–erase endurance [9], and nanoamperes programming energy [10], but can also exploit multilevel conductance states [4]. As a result, it can act as a nonvolatile memory and realize in-memory calculation where data can be processed and stored simultaneously. Thus, the memristor-based system can effectively overcome the obstacles of traditional computing architectures with memory wall, which is of relevance to current and future computing needs, such as, cognitive processing, big-data analysis, reservoir computing, and edge-computing [11].

Privacy preservation is a critical challenge for ANN. The privacy preserving in ANN is to release statistical information from collected data sets without compromising the privacy protection of the individual respondents. An effective method is to introduce a randomized noise mechanism for differential privacy technology to quantify the protection ability. Usually, software-based machine learning algorithms easily generate such randomized noise [2]. However, noisy and distorted data would lead to a degradation of the recognition accuracy in ANN. Accordingly, solutions to balance privacy preserving and recognition accuracy are indeed needed. One popular solution is adopting a specific algorithm, but with considerable computation overhead, which is neither acceptable for a general-purpose computing system such as data

Manuscript received January 30, 2019; revised May 9, 2019; accepted June 8, 2019. Date of publication July 3, 2019; date of current version November 22, 2019. This work was supported by the National Science Foundation under Grant 1855646. *(Corresponding author: Jinhui Wang.)*

J. Fu and Z. Liao are with the Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58102 USA (e-mail: jinyan.fu@ndsu.edu; zhiheng.liao@ndsu.edu).

J. Wang is with the Department of Electrical and Computer Engineering, University of South Alabama, Mobile, AL 36688 USA (e-mail: jwang@southalabama.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2019.2923722

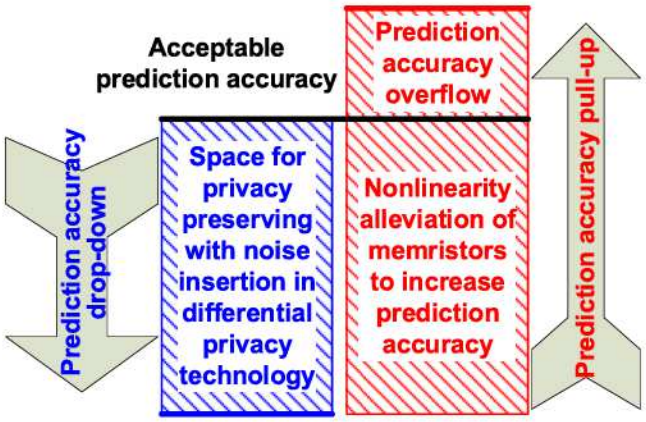

Fig. 1. Concept of memristor-based neuromorphic hardware improvement for privacy-preserving ANN.

center because of the increasing workload, nor sufficient to satisfy the portable and edge computing system due to the resource-constrained, such as mobile devices, wearable devices, and IoT devices. Therefore, in this paper, we propose a memristor-based neuromorphic hardware improvement to enable privacy-preserving ANN without accuracy degradation. That is to use a linear optimization (LO) method to alleviate nonlinearity of memristors for weight updating in ANN to counteract recognition degradation due to noise injection, as shown in Fig. 1. Therefore, privacy-preserving ANN provides enough space for randomizing noisy data to ensure that the publicly visible information do not change much if one individual in the data set changes, which is enabled by differential privacy technology—a strictly provable, quantized, and security-controlled method. Specifically, this paper makes the following contributions.

1) *A method for mitigating the nonlinearity impact on memristor-based privacy-preserving ANN:* To mitigate the impact of memristor's nonlinearity property, an effective, hardware-based LO method with low circuit overhead is proposed, which makes the neuromorphic system become more accurate and applicable for ANN application.
2) *A mechanism enhancing the immunity of memristor-based privacy-preserving ANN to nonlinearity property of memristor device:* By applying the LO method under eight groups of private perturbations that follow the differential privacy theory, the recognition accuracy of ANN is proven to get negligible degradation or even increases than before.
3) *Thorough evaluation:* We evaluate the proposed method on standard image classification tasks [12] and conduct over 1500 simulations that include 4 models, 8 groups of privacy perturbations, and 49 nonlinearity cases.
4) *The tradeoff analysis:* The LO method provides a variety of configuration models, and we discuss how to reduce the cost by selecting the appropriate model while meeting the privacy and accuracy requirements based on the actual device.

The rest of this paper is organized as follows. The related work is presented in Section II. In Section III, the nonlinearity property of memristor and differential privacy are introduced. In Section IV, the approach used to address the nonlinearity is presented. In Section V, the LO method is applied for digits-image recognition with privacy perturbation and is verified through the differentially private transformation algorithm [13] and hardware simulator, NeuroSim+, which is an integrated hardware framework for benchmarking memristors and array architectures [14]. Finally, conclusions are made in Section VI.

II. Related Work

With ANN develops rapidly, it powers intelligent products by extracting patterns and building models. Meantime, data privacy greatly impacts our daily life, such as politics, security, businesses, relationships, health, and finances. The privacy problem is not limited to the threats associated with private data exposures or hacking attempts. It is also possible to glean extra information even if the data are anonymized and the ANN models are inaccessible. Privacy-preserving ANN technologies are proposed to make that ANN transform our society positively without risking our sensitive data, which is mainly conducted by cryptographic approaches or differential privacy approaches [15]. Especially, differential privacy, that is more efficient and popular, resists attacks by adding random noise to the input data, to iterations in a certain algorithm, or to the algorithm output. In 2017, the Google security and privacy team released a Private Aggregation via Teacher Ensembles (PATE) framework [16], which scales to learning tasks with large numbers of output classes and uncurated, imbalanced training data with errors, and it was proven as tighter differential-privacy guarantees. In 2018, the ARDEN framework was proposed to protect the sensitive information via local differentially private and noise training [13]. However, because these technologies are all based on software technologies, the presentence of noise is bound to cause a drop in accuracy and it is impossible to get higher accuracy than without the noise injection. Also, software-based noise injection causes latency problems and computational overhead. Recently, designers proposed to exploit inherent noise with the equivalent error-prone hardware to replace software-based noise to save much power [17], which indicates the hardware can provide a more effective solution to realize privacy-preserving ANN.

In this paper, instead of using traditional software method, we propose a LO method that applies in memristor-based ANN hardware system to improve privacy preserving space for differential privacy technology. The proposed method focuses on mitigating the nonlinearity problem of memristors to enable privacy preserving.

III. Preliminaries

A. Differential Privacy

Differential privacy promises a powerful standard for privacy guarantees either on algorithms or databases [2], [13], [18]. The definition of ε-differential privacy and equation are given below.

Definition [20]: A randomized mechanism A satisfies ε-differential privacy when any adjacent input d and d', and

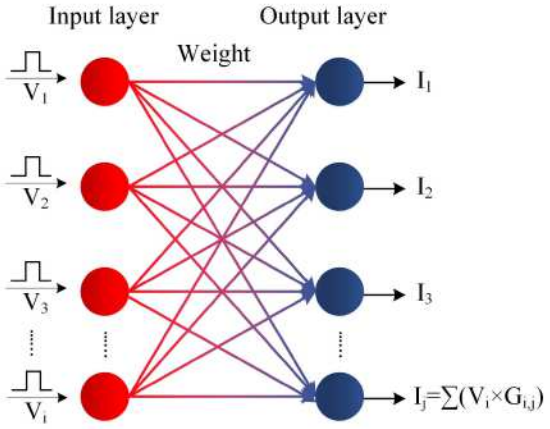

Fig. 2. Neural network between two layers. V_i, $G_{i,j}$, and I_j represent the input signal in ith neuron, the weight of the synapses in jth neuron of output layer and ith neuron in input layer, and the output sum that represent the dot product result of V and G, respectively.

any output S of A hold that

$$\Pr[A(d) = S] \leq e^{\varepsilon} \cdot \Pr[A(d') = S] \tag{1}$$

where d and d' are adjacent inputs that differ in a single entry. In our study, for instance, each training data set is a set of image-label pairs. The d and d' are two sets that only one image-label pair is present in one set and absent in the other. The parameter ε is the privacy budget, which evaluates the privacy guarantee of the randomized mechanism A. A smaller value of ε means the closer recognition accuracies can be gotten from adjacent inputs and indicates a stronger privacy guarantee. By this definition, privacy preservation can be calculated and evaluated through ε.

Typically, adding noise calibrated to the global sensitivity is a general method for approximating a function f, denoted as Δf, which is the maximal value of $||f(d) - f(d')||$ among any input pair of d and d' [18], [21]. For instance, the Laplacian mechanism is defined by

$$Af(d) = f(d) + \text{Lap}(\Delta f/\varepsilon) \tag{2}$$

where Lap $(\Delta f/\varepsilon)$ is a random variable sampled from the Laplace distribution with scale $\Delta f/\varepsilon$.

B. Artificial Neural Networks Hardware

ANN is computing systems vaguely inspired by the biological neural networks that transform inputs to desired outputs by feed-forward networks. As shown in Fig. 2, each neuron in the network takes a weighted sum of the outputs of the prior layer, and then transfers the sum to the next layer. In the hardware implementation, the neural network can be directly mapped into a crossbar from where the inputs are connected into the rows and the outputs are connected into the columns. Memristor-based crossbar circuit can store the synaptic weight and calculate the desired result (sum of product), at the same time, extremely improving the system efficiency. The desirable properties of memristors support the memristor-based crossbar circuit to be a promising substitute technology to traditional

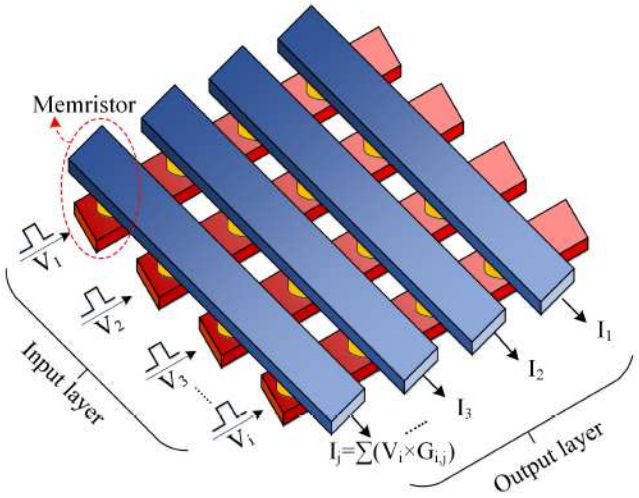

Fig. 3. Hardware implementation of neural networks using memristor crossbar. V_i, $G_{i,j}$, and I_j represent the input signal in ith row, the conductance of the memristor in jth column and ith row, and the output current that represent the dot product result of V and G, respectively.

ones so that researchers begin realizing device-engineering and array-integration hardware implementation of memristors. Usually, as shown in Fig. 3, in the hardware application of the neural network, memristors act as synapses in crossbar structure and locate in each cross point. However, the nonlinear property of the memristor degrades the performance of ANN.

C. Nonlinear Property of Memristors

The conductance of the memristor (G) represents the weight of the synapse and it needs to be updated frequently during the data training process as determined by learning algorithms. In such an updating process, the conductance can either increase in a process as long-term potentiation (LTP) or decrease in a process as long-term depression (LTD), as shown in Fig. 4(a). Ideally, when LTP or LTD occurs, the change in the conductance of an ideal synapse device is proportional to the number of input pulses. Unfortunately, in reality, such change mismatches the input pulse due to the nonlinearity of memristors. For instance, as shown in Fig. 4(b), the curve (black) represents the conductance of an actual memristor device as a function of the number of input pulses where the pulses have the same duration and the same amplitude. While the line (red) in Fig. 4(b) represents the function of the ideal case. In LTP, as shown in Fig. 4(b), assuming in a weight update process, the device's conductance needs to be updated from point a to b. Usually, the corresponding number of pulses is calculated according to the ideal case (red). But, when these pulses are applied to the actual device, instead of changing from point a to b, the device conductance changes from point a to c. Therefore, the actual change of conductance and the required change are quite different. Similar weight updating error occurs in the LTD process. Consequently, the nonlinearity of the memristor causes the weight change of the synapse device to be inconsistent with the change required

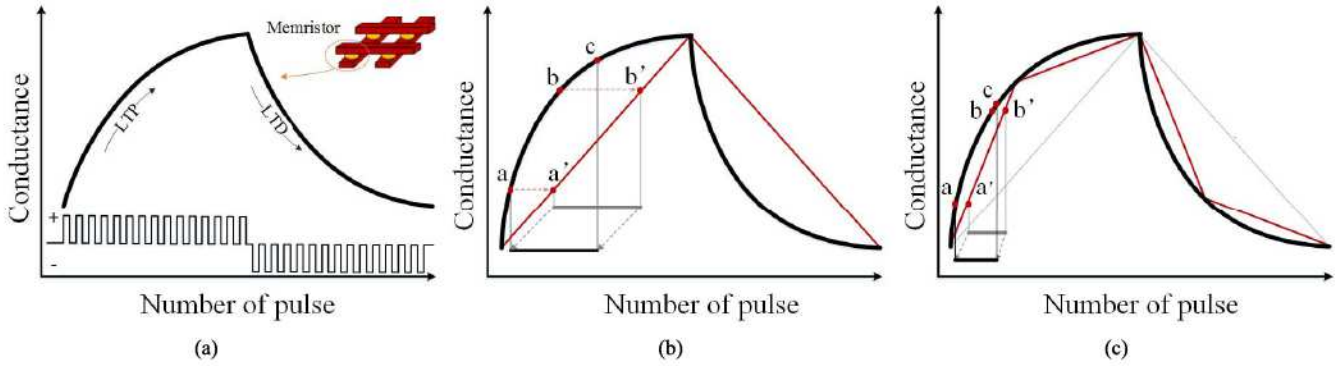

Fig. 4. Conductance change (weight updating) curve. (a) Conductance changes with identical input pulses. (b) Weight updating process based on linear line. (c) Weight updating process based on a piecewise line.

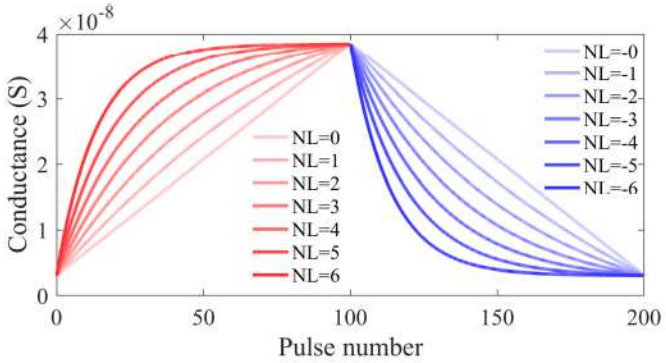

Fig. 5. Conductance change curves with pulse number under various nonlinearity of LTP and LTD.

by the learning algorithm, thereby, reducing the accuracy of ANN's recognition.

Since the nonlinearity property of memristor makes the memristor-based ANN be a challenge especially with the privacy-preservation mechanism introduced, this paper focus on mitigating the impact resulted from the nonlinearity property. We adopt a general conductance change behavior model [22] that is defined by the following equations:

$$G_{\mathrm{LTP}} = B\left(1 - e^{\left(-\frac{P}{A}\right)}\right) + G_{\mathrm{min}} \tag{3}$$

$$G_{\mathrm{LTD}} = -B\left(1 - e^{\left(-\frac{P-P_{\mathrm{max}}}{A}\right)}\right) + G_{\mathrm{max}} \tag{4}$$

$$B = \frac{G_{\mathrm{max}} - \mathrm{G}_{\mathrm{min}}}{1 - e^{\left(\frac{-P_{\mathrm{max}}}{A}\right)}} \tag{5}$$

where G_{max}, G_{min}, and P_{max} are directly extracted from the actual test data [22], which represents the maximum conductance, minimum conductance, and the maximum pulse number required to switch the device between the minimum and maximum conductance states. A and B are the parameter that controls the nonlinear behavior of the weight update. In this model, by adjusting A, the conductance curve is labeled with a nonlinearity value (NL) from $+6$ to -6, which represents the extent to the curve deviates from the ideal linear device and is shown in Fig. 5. Here the positive $(+)$ and negative $(-)$ signs are merely to label LTP and LTD, respectively.

IV. Methodology

A. Linear Optimization Method

To mitigate the impact of the nonlinear property of memristors on privacy-preserving ANN, a LO method is proposed in this paper.

The LO method makes conductance update along a piecewise line instead of an ideal line by making use of multiple pulses with the same amplitude but various duration. As shown in Fig. 4(c), instead of calculating the number of required pulses along the ideal line (gray), the LO method performs the calculation along a polyline (red), which fits the actual device property (black) better. Thus, the error that incurs from the nonlinearity of memristors can be much reduced. For example, we assume that Fig. 4(c) shows the same conductance change curve as depicted in Fig. 4(b), but Fig. 4(c) calculates the needed pulse along a polyline. After the pulses are applied to points a, the conductance difference between the actual points b and c shown in Fig. 4(c) is much smaller than the difference between points b and c shown in Fig. 4(b). In Fig. 4(c), the polyline is composed of two lines and we call it a two-segment LO model. Similarly, based on the number of lines in the polyline, this method can be applied in a three-segment or a four-segment LO model, as shown in Fig. 6.

To implement the ANN with the LO method, first, because of the presence of variations, we need to find the average normalized conductance change curve of over 1000 representative memristor samples. In this case, although the proposed method solves nonlinear problems in varying degrees for each device, it can still greatly alleviate the overall weight-deviations problem of a memristor-based array, which will be discussed in Section V-F. Then, we need to choose split points that can divide the conductance curve into several segments. Thus, the piecewise line can be gotten. Next, the duration of input pulses can be calculated by the slope of the original ideal line, k_0, and the slopes of the piecewise line, $\{k_1, k_2, \ldots, k_n\}$. Specifically, the LO method scales the duration of pulses to k_0/k_i times of the original duration in order to balance the conductance change caused by the nonlinear effect. To implement the LO method, a set of memory, $\log_2(n)$-bit memory, is also needed to store the segment information, which is used to select a slope before each weight update. Here, n represents

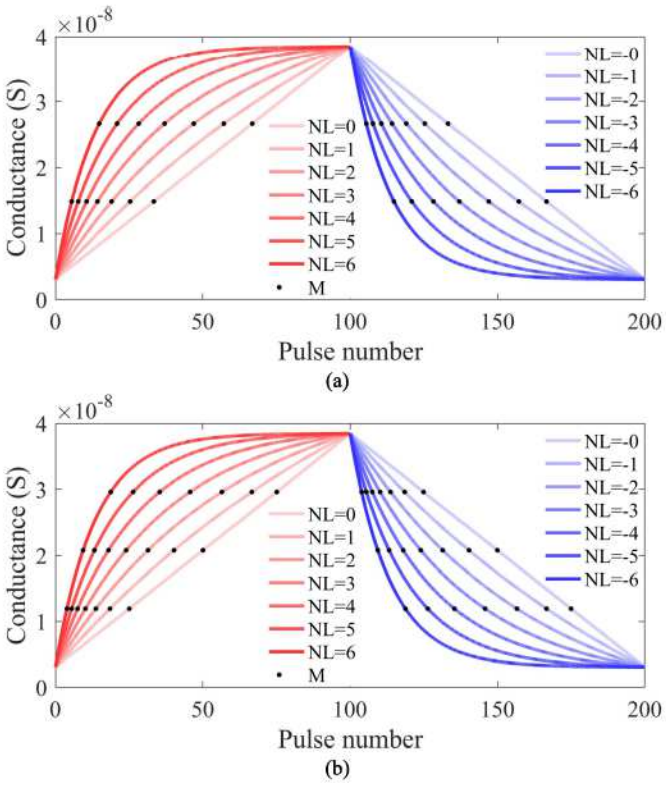

Fig. 6. Segment models and split point M. (a) Three-segment model. (b) Four-segment model.

the segment number, such as two, three, and four. In this way, the larger/smaller the line's slope is, the shorter/longer the duration can be selected. Finally, after each weight update, the comparison operation should be completed to make sure the memory is updated based on the current conductance range of each memristor. This comparison does not need to read precise conductance of memristors but needs to compare with a reference value to recognize the current segment information of memristors.

B. Differentially Private Transformation

According to [18], differential privacy is immune to post-processing: A data analyst, without additional knowledge about the private database, cannot compute a function of the output of a private algorithm and make it less differentially private. This property of differential privacy supports the differentially private transformation algorithm [13], [20]. This paper adopts the Laplacian mechanism and follows the private transformation algorithm [13] where the privacy budget ε is proven to be calculated with given input data and the neural network by the following equation:

$$\varepsilon = 2\sigma \tag{6}$$

where σ is the noise scale in private transformation algorithm in [13].

V. Experimental Evaluation

In this section, digit recognition tasks are used as experimental examples to evaluate the effectiveness of the LO method that aims to mitigate the effect of a memristor's nonlinearity in a hardware implementation. A comprehensive suite of simulations has been conducted to explore the space of privacy-preservation using the proposed LO method in memristor-based ANN. We adopt the neural network hardware platform NeuroSim+ [14] with nonlinearity property injection, as well as private transformation algorithm [13] to perform hardware-based privacy-preserving recognition through the Modified National Institute of Standards and Technology (MNIST) database [12]. The neural network of this simulator includes 400 neurons as input, 100 neurons as a hidden layer, and 10 neurons as an output layer, which is used for recognizing 10 number digits. Each simulation trains up to 125 epochs. Each epoch selects 8000 images randomly from 60 000 training images and takes 10 000 images as a testing data set.

A. LO Models

The proposed LO method needs to select the split points in order to determine the types of pulses. According to the discussion about the weight update error in Section IV, the more the segments, the less the weight update deviation that is caused by the nonlinearity, but the higher the circuit cost. Thus, in order to investigate the tradeoff between the recognition accuracy of privacy-preserving ANN and the cost of the LO method, we conduct three models including two-segment, three-segment, and four-segment models.

As shown in Fig. 6, for two-segment, three-segment, and four-segment models, split points are selected where they can divide the conductance range into two, three, and four equal parts, respectively.

B. Impact of Private Perturbation

To explore the impact of LO method on the memristor-based privacy-preserving ANN, simulations with ε from 4 to 16 that reflects the strength of private preservation are conducted. The smaller the ε is, the larger the noise injection is needed and vice versa. To compare ANN with different nonlinearity of memristor, first, we conduct six groups of simulations with six different memristors that have the same absolute nonlinear value of LTP and LTD process. As shown in Fig. 7, in each figure, the accuracy increases as the ε increases and all cases with the LO method have better performance as compared with the original case without the LO method applied. Among the three models of the LO method, the model with a higher segment shows higher accuracy. What is more, as the memristors' nonlinearity of ANN increase from (1/−1) to (6/−6), the differences between different models become larger and larger.

It concludes that the accuracy of four-segment model not only keeps an accuracy over 70% when ε is larger than 5, but also has less than 10% accuracy difference from NL (1/−1) to NL (6/−6). However, the original model without the LO method gets much more increasing accuracy loss (at least 10%) as nonlinearity increases. Therefore, the four-segment LO method shows more benefits as the nonlinearity of memristor increases in ANN. Furthermore, in some cases,

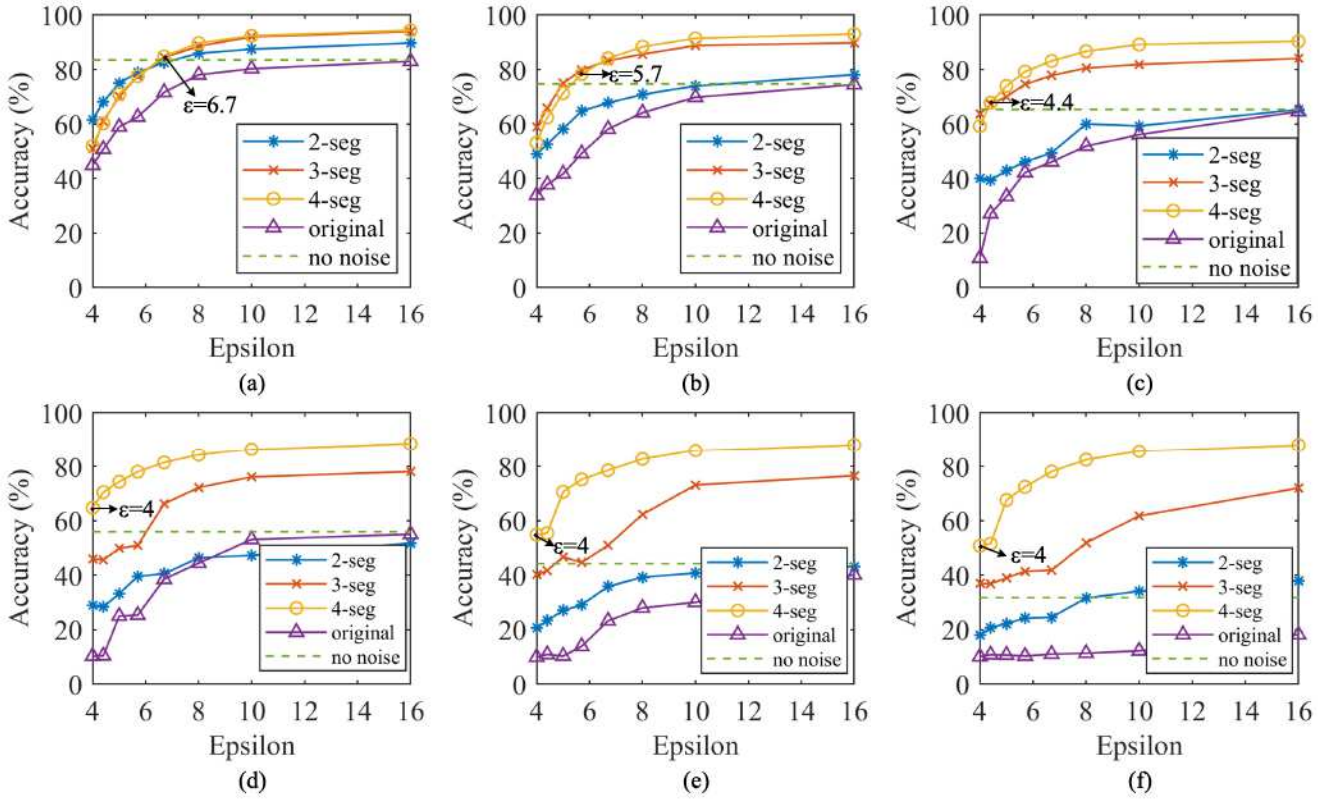

Fig. 7. Recognition accuracy of the MNIST handwriting digits applying four method models with different private perturbation. The NL (x/$-y$) means the LTP nonlinearity of memristor is x and the LTD nonlinearity of memristor is y. (a) NL (1/−1). (b) NL (2/−2). (c) NL (3/−3). (d) NL (4/−4). (e) NL (5/−5). (f) NL (6/−6).

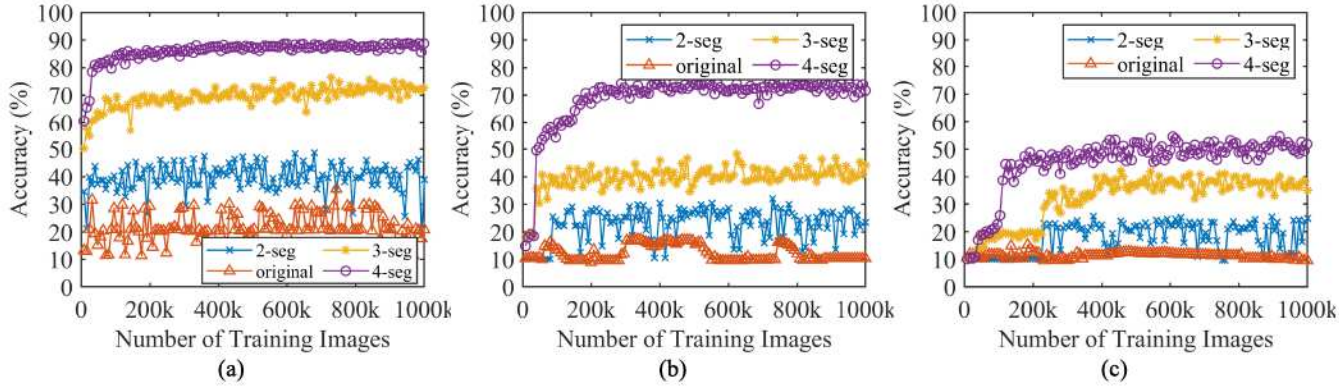

Fig. 8. Recognition accuracy of the MNIST handwriting digits with the training images when the nonlinearity is NL (6/−6) that is considered as the worst nonlinearity case. (a) small noise ($\varepsilon = 16$). (b) middle noise ($\varepsilon = 5.7$). (c) large noise ($\varepsilon = 4$).

the accuracy with the privacy preservation gets even higher accuracy when the LO method applied than the case without privacy preservation, for example, for four-segment model, in NL (1/−1), (2/−2), (3/−3), (4/−4), (5/−5), and (6/−6), when $\varepsilon \geq 6.7, 5.7, 4.4, 4, 4$, and 4, respectively. This indicates the LO method makes the memristor-based ANN hardware get more space for privacy preservation, which would lead to stronger privacy preservation. In addition, the LO method provides various LO models that can be chosen according to the nonlinearity of memristor. When the nonlinearity of the memristor device is relatively small, such as (1/−1) and (2/−2), the results of the three-segment model are similar to that of the four-segment model. Therefore, in this case, considering tradeoff the performance and the cost, the three-segment is a better choice.

C. Stability of LO

Later, we take the worst nonlinearity case of memristor [NL (6/−6)] as an example to explore the impact of different LO method models with $\varepsilon = 16$, 5.7, and 4, on recognition accuracy during the training process. As shown in Fig. 8, the accuracy of each LO model has a short time of fluctuation before convergence, which is decided by the stochastic gradient descent (SGD) algorithm of the hardware simulator. However, these fluctuations are different for different models.

The four-segment model always gets the highest accuracy as well as the most stable accuracy plateau and the three-segment model is also more stable than the two-segment model. The fluctuation of accuracy increases as the noise level increases. In addition, because a larger segment model makes weight updating more fit to the real memristor, causing a smaller weight updating deviation, the case with more segments has higher and more stable recognition accuracy during the training process.

D. Privacy-Preserving Space for Various Nonlinearity

Since for a memristor, the NL of LTP is not usually equal to the NL of LTD [23]–[26], we further investigate 49 different memristor devices whose NLs of LTP and LTD are from 0 to 6 (−6). For those 49 memristors, we perform 49 simulations in four models (original case without the LO method, two, three, and four segment models) and with eight privacy protection noises, respectively. We also simulated the ANN without privacy protection and LO method, resulting in Fig. 9, where the recognition accuracy result is represented by a colored square and the average accuracy of these 49 memristor nonlinearity cases is 55.97%, which is regarded as a parameter that reflects the overall performance of the original ANN. Next, Fig. 10 shows recognition accuracies of ANN hardware without LO applied and with three LO models applied, respectively. It shows the results of three noise levels applied [ε = 16 (small), 5.7 (middle), and 4 (large)]. To compare the overall effectiveness of the proposed method, the average accuracy results are calculated by averaging the accuracy of 49 NL cases, which is shown under each figure. The results show an increasing accuracy, when the segment in LO model increases, the noise level decreases, and the nonlinearity decreases, respectively. When the four-segment LO model is applied and ε equals 5.7, the average can reach 75.46% that is a 47.74% improvement compared to the result 27.72% for that without the LO model. Fig. 11 shows the accuracy improvement in three noise level applying the four-segment model. It shows that the average accuracy improvements are 37.05%, 47.73%, and 34.22%, respectively. Moreover, to study the average accuracy improvement of the LO method, Table I lists the average accuracy for four models with the ε between 4 and 16. As listed in Table I, when the LO method applied, the case with the average accuracy that is larger than the case without noise (55.97%), gives more space for privacy-preservation.

E. Cost Analysis

Models with more segments have better performance of ANN, but they need more cost for storage space and circuits to generate more types of pulses as listed in Table I. However, some nonlinearity cases, their performance with two-segment or three-segment is similar to four-segment model, such as memristors, with NL (1/−1) in Fig. 7(a). Therefore, although the four-segment model achieves the highest average accuracy for 49 NL cases, we should consider different NL cases independently to lower the unnecessary cost of storage space and circuits. Also, the LO method brings additional access in order to read and write the memory used for storing segment information. As shown in Fig. 12, these accesses do not induce extra latency, because they are conducted with memristors' reading and calculation at the same time, nearly hidden and covered in memristors' operations. The proposed method only brings latency overhead for the comparison operation, as shown in Fig. 12. But such comparison time is only a small ratio in one cycle. Finally, as compared with state-of-the-art works [13], [22], [27]–[30], the proposed method needs less circuit complexity including a simpler pulse generator for the same amplitude pluses.

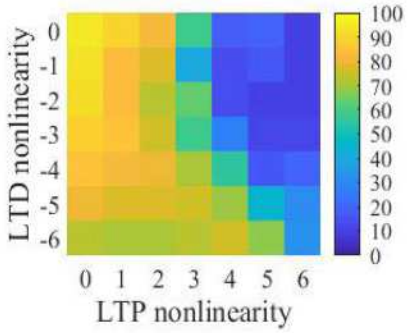

Fig. 9. Recognition accuracy of the MNIST handwriting digits without the privacy-preservation and LO method. The average accuracy of the 49 memristors' nonlinearity cases is 55.97%.

TABLE I
COST AND AVERAGE ACCURACY WITH DIFFERENT MODELS

Model		original	2-seg	3-seg	4-seg
Bit # cost		0	1	2	2
Pulses # cost		2	4	6	8
Average Accuracy (%) for 49 NL cases	ε=4	19.72%	35.23%	47.21%	53.94%
	ε=4.4	20.70%	39.33%	53.66%	**60.94%**
	ε=5	24.37%	45.02%	**58.62%**	**68.87%**
	ε=5.7	27.72%	50.06%	**64.77%**	**75.46%**
	ε=6.7	33.56%	55.20%	**71.95%**	**81.17%**
	ε=8	40.93%	**62.73%**	**78.31%**	**85.94%**
	ε=10	48.07%	**65.59%**	**82.69%**	**89.18%**
	ε=16	54.25%	**69.56%**	**85.59%**	**91.29%**
	No noise	55.97%	-	-	-

[a] The data in bold show the cases that have higher accuracy than the original case without LO method applied (55.97%) as well as without privacy preservation.

F. Variations of Memristors

The physical mechanism of the conductance modulation in most prospective synaptic devices is typically an ionic reconfiguration process based on electro/thermodynamics. This thermally activated ion migration and process variations are responsible for unavoidable variations including nonlinearity, device-to-device, cycle-to-cycle, and ON/OFF conductance variations [22], [31], [32]. Considering variations exist among real devices, we simulate five variations that subject to standard normal distribution $N(\mu, \sigma)$ to explore the effectiveness of the four-segment LO method. In our simulation, minimum conductance, maximum conductance, and device-to-device

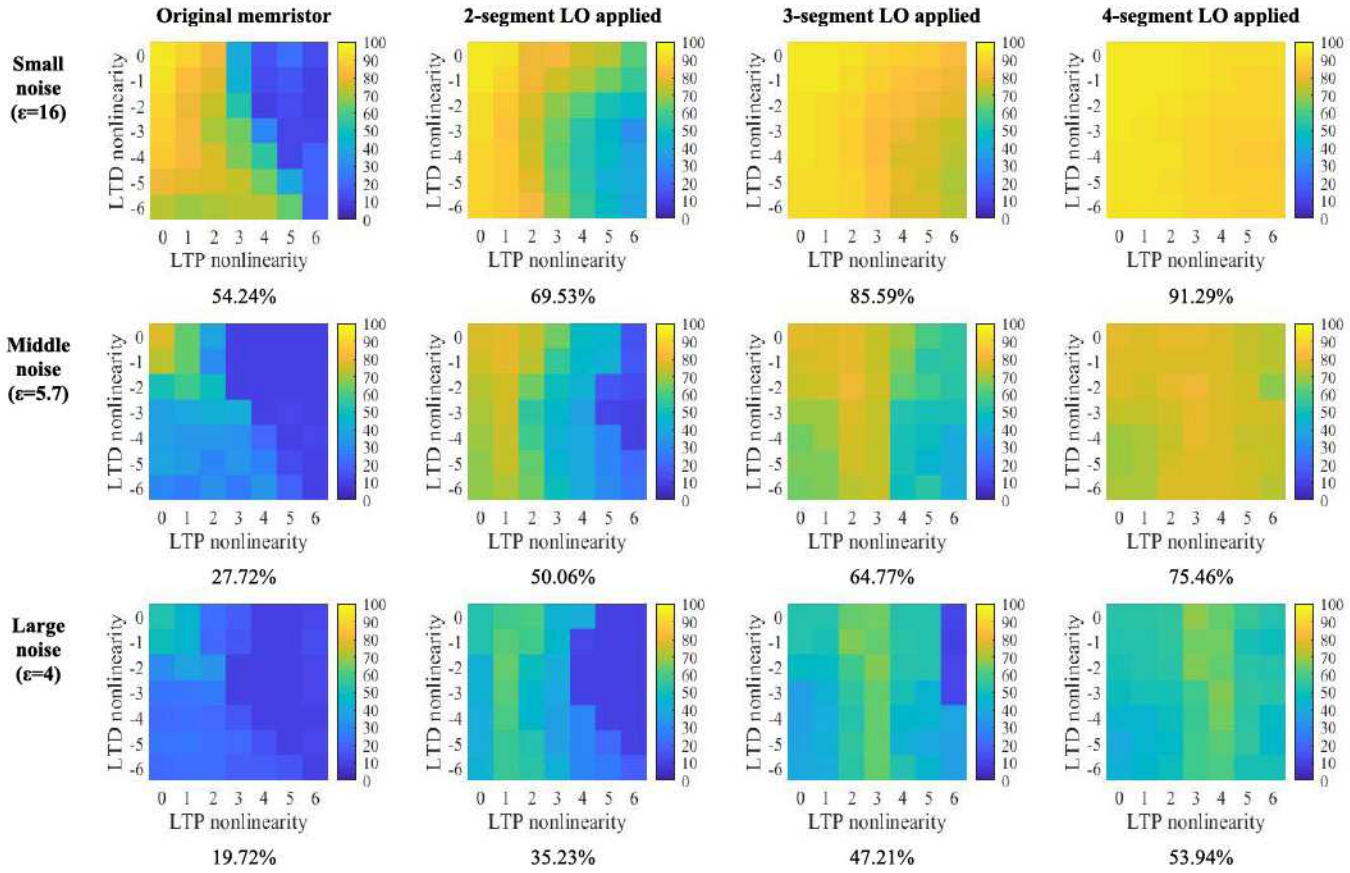

Fig. 10. Recognition accuracy of memristor-based ANN with different models for 49 nonlinearity cases of memristor.

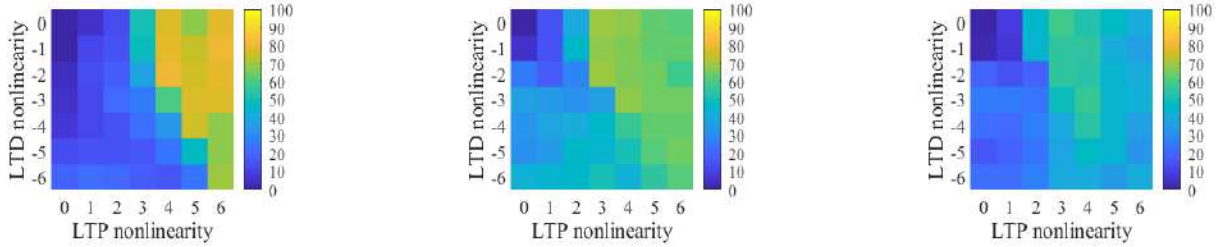

Fig. 11. Recognition accuracy improvement of four-segment LO model. The average accuracy improvement of each figure is shown under each figure. (a) small noise ($\varepsilon = 16$). (b) middle noise ($\varepsilon = 5.7$). (c) large noise ($\varepsilon = 4$).

variation subject to N ($G_{\max}$, $\sigma \times G_{\max}$), N ($G_{\min}$, $\sigma \times G_{\min}$), and N [NL(LTP), σ] and N [NL(LTD), σ], respectively. Cycle-to-cycle variation is illustrated in the following equation [22]:

$$G = G + (G_{\max} - G_{\min}) \times N(0, \sigma) \times (N_{\mathrm{p}})^{\alpha} \tag{7}$$

where Np represents the needed pulse number in each weight update, α represents the impact of Np and it is set to 0.5 in our simulations. ON/OFF ratios are configured as 16 in variation 1 and 14 in variation 2. For variations 1 and 2 in Table II, we set σ of minimum conductance, maximum conductance, device-to-device, and cycle-to-cycle variation as 6%, 6%, 1/1, 1%, and 18%, 18%, 3/3, 3%, respectively [33].

TABLE II
RECOGNITION ACCURACY WITH DIFFERENT VARIATIONS

Epsilon		16		5.7		4	
Method		LO	Original	LO	Original	LO	Original
Variation 1	(2/-2)	88.89%	10.98%	76.75%	11.02%	61.59%	11.51%
	(4/-4)	84.12%	10.99%	69.27%	10.81%	50.56%	11.39%
	(6/-6)	76.57%	11.36%	47.45%	10.90%	36.60%	9.53%
Variation 2	(2/-2)	73.22%	10.44%	39.63%	10.35%	37.63%	13.27%
	(4/-4)	64.15%	11.01%	36.94%	10.74%	29.76%	12.17%
	(6/-6)	58.09%	11.73%	38.80%	11.51%	29.43%	10.11%

In the presence of variations, the proposed method can keep recognition accuracies higher than 75% when the noise level and variations are both small. Under the same circumstance

TABLE III
COMPARISON OF THE STATE-OF-ART

	[13] in 2018	[22] in 2015	[27] in 2016	[28] in 2018	[29] in 2018	[30] in 2018	This work
Without precise read-before-write	√	√	√	×	√	√	√
Without always change pulse amplitude	√	√	√	√	×	×	√
Without always change pulse duration	√	×	√	×	√	√	√
Almost immune to nonlinearity	×	√	×	√	√	√	√
No need for algorithm optimization	×	√	√	√	√	√	√

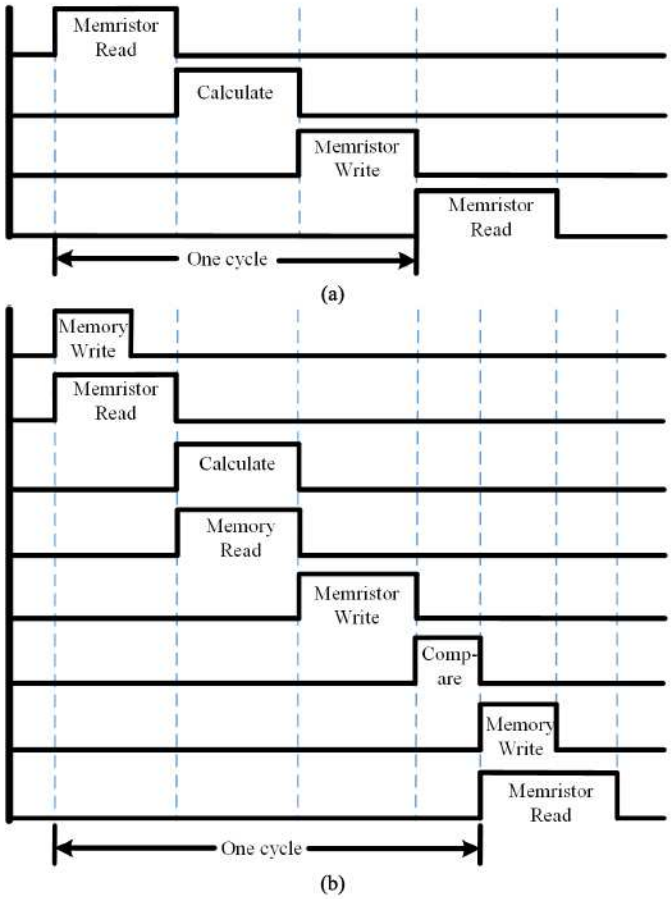

Fig. 12. One cycle timing schematic in weight update process. (a) Without LO method. (b) With LO method. The memory represents the added memory component and the memristor represents the memristor that acts as a synapse in neural network.

without the LO method, recognition accuracies are not higher than 12%. Although as the variations and noise level increase, the accuracies of cases with the LO method decrease, they are still much higher than the accuracies without the LO method. It concludes that in real device condition when various variations exist, the proposed LO method is still proven to be an effective method for privacy preserving.

G. Comparison With Other Works

As listed in Table III, instead of optimizing the algorithm for privacy preservation [13], the proposed LO method is simple and feasible to addresses accuracy degradation due to privacy preservation by optimizing ANN in the hardware. As compared with the state-of-the-art [13], [22], [27]–[30], the LO method does not need to read the precise conductance of memristor before every writing operation and does not need to change the amplitude of the update-pulses each time so that it avoids complex peripheral circuits. What is more, with a four-segment model, the LO method is almost immune to the nonlinearity of memristor. Because our simulations are all based on the standard SGD algorithm and a regular hardware simulator, the recognition results still have a large space to be improved by using a more efficient ANN algorithm or by high-performance memristor devices. The method we propose is a universal method that works for all memristor-based hardware with nonlinear characteristics. Accordingly, the LO method is an effective technique to address the weight deviation issue caused by the nonlinearity property of memristors in privacy-preserving ANN; also, it provides multiple configurations to meet different requirements of privacy preservation.

VI. CONCLUSION

In this paper, the LO method is proposed to improve the performance of memristor-based privacy-preserving ANN and it is verified based on the MNIST database, the differentially private algorithm, and the memristor-based neural network simulator. Instead of adopting the traditional algorithm-based technology, the LO method focuses on hardware implementation to enable privacy-preserving ANN. It does not need to read the precise conductance of memristor before every writing operation in weight-updating process and does not need to change the amplitude of the update-pulses each time in ANN, which avoids complex peripheral circuits. The two-segment, three-segment, and four-segment models for 49 types of memristors with nonlinearity from (0/−0) to (6/−6) have been developed to investigate the effectiveness of the proposed method, the results indicate 34.22%–47.73% average recognition accuracy improvement when the privacy budget ε ranges from 4 to 16. It concludes: 1) the proposed privacy-preserving ANN has an increasing accuracy, when the segment in the LO model increases, the noise level decreases, and the nonlinearity decreases, respectively; 2) a LO model with more segments not only has a stronger immunity to nonlinearity but also gets higher and more stable accuracy; and 3) the

proposed method is proven to be effective when variations exist. Furthermore, in some cases, since the accuracy with privacy preservation gets even higher accuracy, the LO method is applied to provide more space and margin for privacy preservation. Finally, the LO method aims at mitigating the nonlinearity impact of memristor devices; therefore, it can be adapted to many other memristor-based hardware systems. Consequently, the LO method is proven to be an effective technique that can prevent accuracy loss and increase privacy preservation space for privacy-preserving ANN.

REFERENCES

[1] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[2] M. Abadi *et al.*, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2016, pp. 308–318.

[3] *Security and Privacy Issues with Internet of Things (IoT)*. Accessed: May 1, 2019. [Online]. Available: https://theipcenter.com/2018/11/security-and-privacy-issues-with-internet-of-things-iot/

[4] L. O. Chua, "Memristor-the missing circuit element," *IEEE Trans. Circuit Theory*, vol. CT-18, no. 5, pp. 507–519, Sep. 1971.

[5] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, p. 80, 2008.

[6] B. Govoreanu *et al.*, "10×10 nm^2 Hf/HfOx crossbar resistive RAM with excellent performance, reliability and low-energy operation," in *IEDM Tech. Dig.*, Dec. 2011, pp. 31.6.1–31.6.4.

[7] A. C. Torrezan, J. P. Strachan, G. Medeiros-Ribeiro, and R. S. Williams, "Sub-nanosecond switching of a tantalum oxide memristor," *Nanotechnology*, vol. 22, no. 48, 2011, Art. no. 485203.

[8] B. J. Choi *et al.*, "High-speed and low-energy nitride memristors," *Adv. Funct. Mater.*, vol. 26, no. 29, pp. 5290–5296, Aug. 2016.

[9] K.-H. Kim, S. H. Jo, S. Gaba, and W. Lu, "Nanoscale resistive memory with intrinsic diode characteristics and long endurance," *Appl. Phys. Lett.*, vol. 96, no. 5, 2010, Art. no. 053106.

[10] J. Zhou, F. Cai, Q. Wang, B. Chen, S. Gaba, and W. D. Lu, "Very low-programming-current RRAM with self-rectifying characteristics," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 404–407, Apr. 2016.

[11] M. A. Zidan, J. P. Strachan, and W. D. Lu, "The future of electronics based on memristive systems," *Nature Electron.*, vol. 1, no. 1, p. 22, Jan. 2018.

[12] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[13] J. Wang, J. Zhang, W. Bao, X. Zhu, B. Cao, and P. S. Yu, "Not just privacy: Improving performance of private deep learning in mobile cloud," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2018, pp. 2407–2416.

[14] P. Y. Chen, X. Peng, and S. Yu, "NeuroSim+: An integrated device-to-algorithm framework for benchmarking synaptic devices and array architectures," in *IEDM Tech. Dig.*, Dec. 2017, pp. 1–6.

[15] M. Al-Rubaie and J. M. Chang, "Privacy-preserving machine learning: Threats and solutions," *IEEE Security Privacy*, vol. 17, no. 2, pp. 49–58, Mar./Apr. 2019.

[16] N. Papernot, S. Song, I. Mironov, A. Raghunathan, K. Talwar, and Ú. Erlingsson, "Scalable private learning with PATE," 2018, *arXiv:1802.08908*. [Online]. Available: https://arxiv.org/abs/1802.08908

[17] L. Yang and B. Murmann, "Approximate SRAM for energy-efficient, privacy-preserving convolutional neural networks," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, Jul. 2017, pp. 689–694.

[18] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014.

[19] A. Beimel, H. Brenner, S. P. Kasiviswanathan, and K. Nissim, "Bounds on the sample complexity for private learning and private data release," *Mach. Learn.*, vol. 94, no. 3, pp. 401–437, Mar. 2014.

[20] C. Dwork, "Differential privacy," *Encyclopedia of Cryptography and Security*, H. C. A. van Tilborg and S. Jajodia, Eds. Boston, MA, USA: Springer, 2011, pp. 338–340.

[21] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Theory Cryptography Conf.*, Mar. 2006, pp. 265–284.

[22] P.-Y. Chen *et al.*, "Mitigating effects of non-ideal synaptic device characteristics for on-chip learning," in *Proc. IEEE/ACM IICCAD*, Nov. 2015, pp. 194–199.

[23] J. Woo *et al.*, "Improved synaptic behavior under identical pulses using AlO_x/HfO_2 bilayer RRAM array for neuromorphic systems," *IEEE Electron Device Lett.*, vol. 37, no. 8, pp. 994–997, Aug. 2016.

[24] S. Park *et al.*, "Neuromorphic speech systems using advanced ReRAM-based synapse," in *IEDM Tech. Dig.*, Dec.2013, pp. 25.6.1–25.6.4.

[25] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano Lett.*, vol. 10, no. 4, pp. 1297–1301, 2010.

[26] L. Gao *et al.*, "Fully parallel write/read in resistive synaptic array for accelerating on-chip learning," *Nanotechnology*, vol. 26, p. 455204, Nov. 2015.

[27] I.-T. Wang, C.-C. Chang, L.-W. Chiu, T. Chou, and T.-H. Hou, "3D Ta/TaO_x/TiO_2/Ti synaptic array and linearity tuning of weight update for hardware neural network applications," *Nanotechnology*, vol. 27, no. 36, Aug. 2016, Art. no. 365204.

[28] C. Li *et al.*, "Analogue signal and image processing with large memristor crossbars," *Nature Electron.*, vol. 1, no. 1, p. 52, Jan. 2018.

[29] C. Li *et al.*, "Efficient and self-adaptive *in-situ* learning in multilayer memristor neural networks," *Nat. Commun.*, vol. 9, no. 1, p. 2385, Jun. 2018.

[30] M. Hu *et al.*, "Memristor-based analog computation and neural network classification with a dot product engine," *Adv. Mater.*, vol. 30, no. 9, 2018, Art. no. 1705914.

[31] S. Kim, M. Lim, Y. Kim, H. D. Kim, and S. J. Choi, "Impact of synaptic device variations on pattern recognition accuracy in a hardware neural network," *Sci. Rep.*, vol. 8, no. 1, p. 2638, Feb. 2018.

[32] D. Ielmini, "Modeling the universal set/reset characteristics of bipolar RRAM by field-and temperature-driven filament growth," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4309–4317, Dec. 2011.

[33] P.-Y. Chen, X. Peng, and S. Yu, "NeuroSim: A circuit-level macro model for benchmarking neuro-inspired architectures in online learning," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 37, no. 12, pp. 3067–3080, Dec. 2018.

Jingyan Fu (S'17) received the B.E. and M.S. degrees in electrical engineering from the Beijing University of Technology, Beijing, China, in 2014 and 2017, respectively. She is currently working toward the Ph.D. degree at the North Dakota State University, Fargo, ND, USA.

Her current research interests include hardware design and algorithm optimization for neuromorphic computing.

Zhiheng Liao (S'17) received the B.E. and M.S. degrees in electrical engineering from the Beijing University of Technology, Beijing, China, in 2014 and 2017, respectively. He is currently working toward the Ph.D. degree at the North Dakota State University, Fargo, ND, USA.

His current research interests include emerging device and algorithm optimization for neuromorphic computing.

Jinhui Wang (M'13–SM'19) received the B.E. degree in electrical engineering from Hebei University, Hebei, China, and the Ph.D. degree from the Beijing University of Technology, Beijing, China.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of South Alabama, Mobile, AL, USA. He has authored or coauthored more than 120 publications and 20 patents in the emerging semiconductor technologies. His current research interests include neuromorphic computing, low-power, high-performance, and variation-tolerant IC design, 3-D IC, and thermal solution in VLSI.