Deep Learning Approach to Estimating Work Function Fluctuation of Gate-All-Around Silicon Nanosheet MOSFETs with A Ferroelectric HZO Layer

Rajat Butola[1,2], Yiming Li[1-6,*], and Sekhar Reddy Kola [1,2]

[1]Parallel and Scientific Computing Laboratory; [2]EECS International Graduate Program; [3]Institute of Communications Engineering; [4]Department of Electrical and Computer Engineering, [5]Institute of Biomedical Engineering, [6]Center for mmWave Smart Radar System and Technologies, National Yang-Ming Chiao Tung University, 1001 Ta-Hsueh Rd., Hsinchu 300, Taiwan. *Tel: +886-3-5712121 ext. 52974; Fax: +886-3-5726639; Email: ymli@nycu.edu.tw

2022 6th IEEE Electron Devices Technology & Manufacturing Conference (EDTM) | 978-1-6654-2178-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/EDTM53872.2022.9798172

Abstract

Highly scaled MOSFETs are suffering from various fluctuations. In this paper, an artificial neural network (ANN) device modeling technique is reported for gate-all-around silicon nanosheet MOSFETs (GAA Si NS MOSFETs). The well-trained ANN model can rapidly and accurately estimate the effect of work function fluctuation (WKF) on device characteristic. Our model is generic because it can be successfully evaluated on the device with a ferroelectric HZO layer which have material and structural dissimilarity with the GAA NS device.

(Keywords: deep learning, artificial neural network, work function fluctuation, gate-all-around, nanosheet, ferroelectric layer, HZO, MOSFETs)

I. Introduction

The focus of IC industry is shifted towards changing the structure of transistors to overcome the excessive leakage current problem of planer devices. In that process, the FinFETs have proved to become a solution to control it significantly. But they are challenged severely as the technology nodes are further scaling down to sub-5 nm [1]. The FinFETs are experiencing greater short channel effects (SCE) and, as a result, the performance goals cannot meet. Thus, the device structure is modified into GAA. The GAA devices are emerging as the most promising successors and have better channel control compared to FinFETs. Apart from this, the GAA NS devices also provide a degree of flexibility which FinFETs lack. Besides, the device with a ferroelectric layer exhibit steep switching behavior and have high on current (I_{ON}) which are useful in low power and high speed applications [2]. Generally, the high scaled transistors are suffered from various fluctuations such as WKF, random dopant fluctuation, interface trap fluctuation, and line edge roughness [3]-[5].

After the 4th industrial revolution, applications of machine learning and deep learning are increased in semiconductor manufacturing and device modeling [6-7]. The ANN, one of deep learning techniques, has also been advanced to explore device variability [8]. Notably, in contrast to compact modeling, device modeling based on ANN is promising in communities of device engineering and circuit design [9]. Thus, it is worth further exploring its potentials for advanced nano-device modeling.

In this work, DC characteristics variations of GAA Si NS MOSFETs induced by WKF are explored using an ANN model. The versatility of the proposed model is further examined by evaluating DC characteristics of GAA Si NS MOSFETs with a ferroelectric HZO layer. Additionally, the performance of the model is also assessed by accurately extracting the variability of the threshold voltage (V_{TH}).

II. Device Structure and ANN Modeling

A. Device Structure & Simulation

The 3D schematic illustrations of GAA Si NS MOSFETs with and without a ferroelectric layer used for this study are shown in Figs. 1(a) and (b) respectively. For both devices, the channel doping concentration is empirically 5×10^{17} cm^{-3}. MOSFETs with ferroelectric layers attracted enormous attention of researchers due to better electrical characteristics [10]. A ferroelectric layer $Hf_{0.5}Zr_{0.5}O_2$ (HZO) material. is considered in this study. For WKF, the TiN (Titanium Nitride) layer is used which is composed of 2×2 nm^2 grains. Two different orientations of TiN, <200> and <111>, with 60% and 40% probabilities are used with work functions (WKs) 4.6 and 4.4 eV respectively [3]. The GAA Si NS MOSFETs have metal grain numbers (MGN) 80 and 104 respectively.

To study the DC characteristics including the V_{TH} variability, 1,800 devices with varying the WKF, are created and simulated using our own 3D device simulation. The transfer characteristics (I_D-V_G) are simulated and the magnitudes of V_{TH} are extracted for the devices with and without HZO. Table I lists the key device parameters used in our device simulation.

B. ANN modeling

The calibrated simulation IV characteristic curves are used to train the ML model. The ANN architecture, used in this work, is shown in Fig. 2. The number of hidden layers and neurons in each hidden layer are the hyperparameters of the model which are tuned to achieve maximum accuracy. Moreover, the activation function and optimizer are also selected carefully. The ANN model has one input layer, two fully connected hidden layers and one output layer with "hyperbolic tangent" (tanh) as activation function and "adam" as optimizer. The hidden layers have 45 neurons in each layer. The fitting capability and accuracy of the ANN model are tested for the current application used in this paper.

III. Results and Discussion

The ANN modeling consists of training and testing as two main processes of ML. The model learns the useful hidden information from input data in training phase and then the trained model is tested on unseen data in testing phase. Here, the WKF patterns are used as input features to train the ANN model and the drain current (I_D) for each corresponding gate voltage (V_G) is estimated as output. The training of model is considered done when the loss function reduces to its minimum value. Different set of hypermeters along with different activation functions are tried and tested to achieve the minimum error rate. The loss function versus number of epoch curves for different hyperparameters, which are used in this paper, are plotted in Figs. 3 (a)-(d).

Once, the training is completed for the GAA Si NS MOSFET devices, the ANN model is applied on the testing dataset. The training efficiency of the model is calculated using root mean square error (RMSE) as it is better in terms of reflecting model's performance. The training and testing rmse are evaluated as 1.57×10^{-07} and 2.59×10^{-07}, respectively, which confirms that the model is trained successfully. All the testing curves are shown in Fig. 4(a) and three testing curves for better visualization are shown in 4(b). It can be examined from the results that the ANN outputs (line) have a good match with the target values (symbol).

Next, the generality of the model is examined by exposing it to the GAA Si NS MOSFETs with ferroelectric HZO layer device data without any changes to the ANN model architecture. It is known that both the devices have material and structural dissimilarities between them. This device has additional complexity due to its negative DIBL effect because of ferroelectric material gate. Despite having all these differences, the model is still able to predict the targets values successfully with high accuracy. The resulting test curves for GAA Si NS MOSFETs with HZO are shown in Figs. 4 (c) and (d).

The WKF also causes the variation in output parameters. For this work, one such important parameter, V_{TH} is extracted from the simulated data. To check the capability of estimating the changes in this key output parameter, the proposed ANN model is applied to extract the V_{TH} values from the predicted *I-V* curves. The true V_{TH} and ANN extracted V_{TH} values of the explored devices with and without HZO film are then compared. The true versus predicted plot, for V_{TH} variability of both devices, is shown in Fig. 5. The results proved that the ANN model has the accurate predictions for different V_{TH} values.

IV. Conclusions

The ANN-based modeling has been evaluated for GAA Si NS MOSFETs. The characteristic variability of the explored devices induced by WKF has been modeled using the DL technique. We have explored the generality of the ANN model by evaluating different structures, i.e., devices with HZO film, by reproducing the *I-V* characteristics. The model capability has been further tested by extracting the magnitudes of V_{TH} for two different MOSFET devices. The results show the powerful representation capability of the model. The proposed model is an effective alternative to compact model for modeling the advanced complex emerging devices.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology (MOST), Taiwan, under Grant MOST 110-2221-E-A49-139, Grant MOST 109-2221-E-009-033, and Grant MOST 109-2634-F-009-030, and in part by the "Center for mm Wave Smart Radar Systems and Technologies" under the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan.

References

[1] A. Razavieh, P. Zeitzoff and E. J. Nowak, IEEE Trans. Nanotechnol., 18, p. 999 (2019).
[2] L. Tu et al., Adv. Electron. Mater., 4, p. 1 (2018).
[3] Y. Li, H. Chang, C. Lai, P. Chao and C. Chen, IEDM, p. 34.4.1 (2015),
[4] S. Oh and C. Shin, IEEE Trans. on Electron Devices, 63, p. 4617 (2016).
[5] S. R. Kola, Y. Li and N. Thoti, SISPAD, p. 1 (2020).
[6] T. Moriya, EDTM, p. 1 (2021)
[7] Y. Lei, X. Huo and B. Yan, EDTM, p. 154 (2018).
[8] C. Akbar, Y. Li and W.-L. Sung, IEEE Trans. Semiconductor Manuf., 344, p. 513 (2021).
[9] L. Zhang and M. Chan, J. Comput. Electron, 16, p. 825 (2017).
[10] M. Hoffmann, S. Slesazeck, and T. Mikolajick, APL Mater., 9, p. 020902-1 (2021).

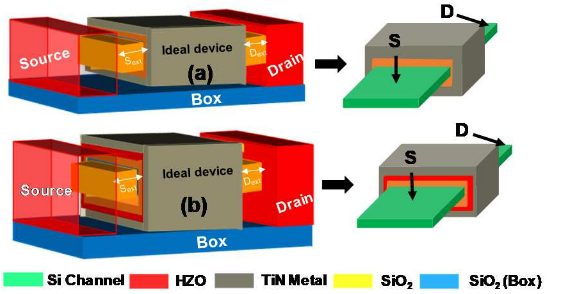

Fig. 1: Schematic structures of (a) GAA Si NS MOSFET (b) GAA Si NS MOSFET with FE material HZO layer. The devices are simulated for obtaining the training and test data.

Table I: List of device parameters used in the device simulation for both structures.

Parameters	Value	Parameters	Value
Gate Length (LG) (nm)	12	S_{ext}/D_{ext} Doping (cm-3)	4.8×10^{18}
Nanosheet Width (nm)	25	HZO Thickness (nm)	4
Nanosheet Height (nm)	5	SiO2 Thickness (nm)	0.6
S_{ext}/D_{ext} Length (nm)	5	TiN <200> WKF (eV)	4.6
S/D Doping (cm-3)	1×10^{20}	TiN <111> WKF (eV)	4.4

Table II: List of hyperparameters with values utilized in the ANN model.

Hyper Parameters	Value
Input Dimension	(900, 80)
Hidden Layers	2
Neurons in Each Hidden Layers	45
Output Dimension	36
Activation Function	tanh
Optimizer	Adam
Loss Function	MSE
Batch Size	5
Epochs	100
Learning Rate	0.001

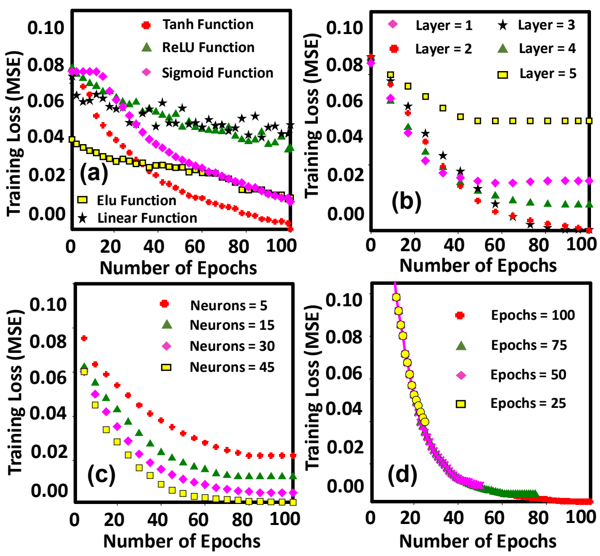

Fig. 3: The loss function versus number of epochs for the different ANN model hyperparameters: (a) activation function, (b) layers (c) neurons (d) epochs.

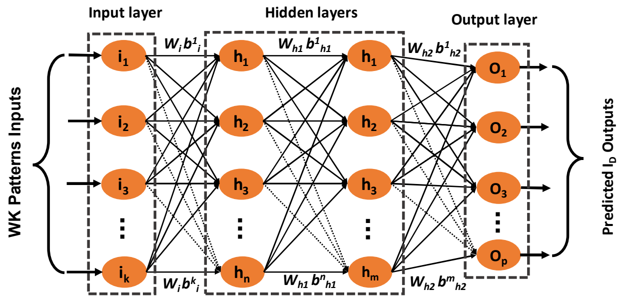

Fig. 2: An illustration of the explored ANN architecture used in this work. It consists of 4 layers: one input layer, 2 hidden layers and one output layer. The WKFs with the MGN = 80 are used as the input features and I_D values corresponding to each V_G are predicted as the outputs.

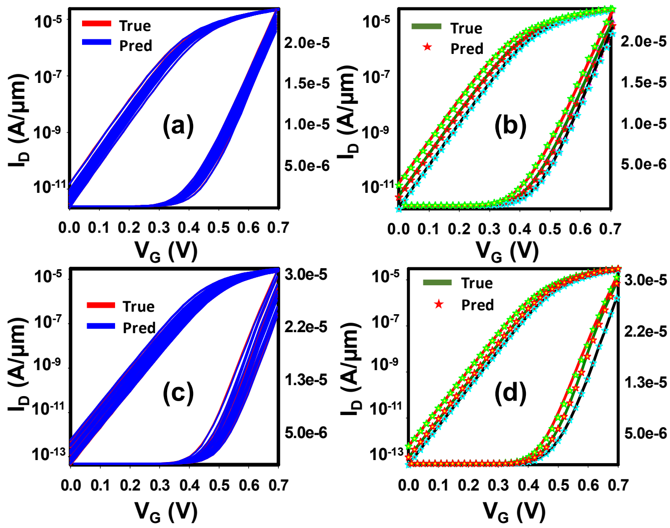

Fig. 4: Transfer characteristic curves of true simulation data and prediction results of the ANN model for both the devices in linear and logarithmic scale. The curves show a good match as they clearly overlap to each other.

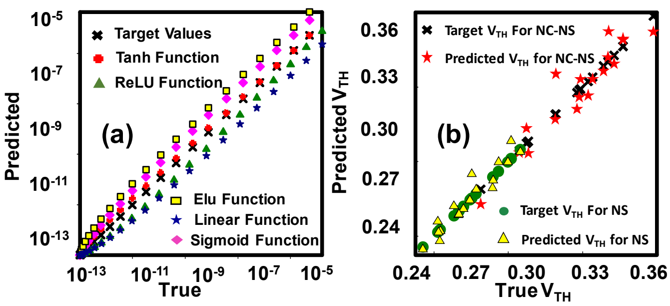

Fig. 5: (a) Comparison between the true and ANN predicted values for different activation functions for 2 hidden layers with 45 neurons in each and 100 epochs. The plot clearly represents that "tanh" function gives the best fitted values for the predicted I_D compared to other activation functions. (b) It shows the comparison between the true and predicted values of variable threshold voltage (V_{TH}) for both the devices. It is evaluated from the plots that the ANN predicted the values for both devices with less than 1% error.