Electrostatic Potential Profile Generator for Two-Dimensional Semiconductor Devices

Seung-Cheol Han, Jonghyun Choi, and Sung-Min Hong
School of EECS, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, email: smhong@gist.ac.kr

***Abstract*—As efficiency is one of the bottlenecks of device simulation, we propose to employ deep neural networks to generate two-dimensional electrostatic potential profiles for efficiency. Supervising with previously obtained simulation results for various BJT devices, we train deep neural networks to generate an electrostatic potential profile as an initial guess for a non-equilibrium condition with estimating carrier densities by the frozen field simulation. With the generated potential profiles, we significantly reduce the number of Newton iterations without loss of accuracy.**

I. Introduction

Deep neural networks are utilized in many fields including computer vision, natural language processing, and expert systems. In semiconductor research, deep neural networks are used in two different aspects. First, there are many efforts to develop a new hardware system to implement efficient deep neural networks [1], [2]. Second, the deep networks are used as an optimization tool in the technology development cycle [3], [4]. To the best of our knowledge, there is hardly an effort to improve efficiency of the device simulation. Since we have tremendous amount of solutions as results of multiple simulations, the deep neural network which learn the existing solutions can give us reasonably good estimation to the real solution. We propose to use deep neural network to quickly analyze the two-dimensional semiconductor devices.

In the device simulation, the computational cost depends on the number of bias points. Even when only the final bias point is of interest, the bias voltages should be gradually increased due to the severe nonlinearity of the semiconductor device equations. In Newton method, good initial solutions will expedite the convergence thus improve computational efficiency significantly. If the generated profiles are close enough to the numerical solutions, the bias ramping process can be skipped; the skip provides further computational gain. In our previous work [5], we showed that the deep neural networks could generate good initial profiles for the one-dimensional structures. As an extension to our previous work, we here propose deep neural networks for two-dimensional structures such as bipolar junction transistors (BJTs).

II. Neural networks

As discussed in our previous work [5], the electrostatic potential is a key quantity to be generated by deep neural networks. Under a fixed electrostatic potential profile, the electron and hole continuity equations become linear or at most locally nonlinear, depending on the adopted physical models. For this reason, a deep neural network is designed to generate electrostatic potential profiles.

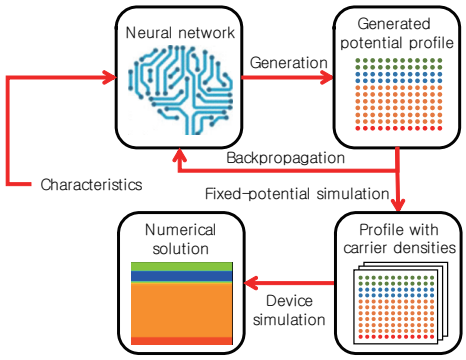


Fig. 1. A diagram for the proposed method. We consider two-dimensional structures in this work.

Fig. 1 shows a diagram for the proposed method. We use some important device parameters (such as bias voltages and doping profiles) as input and use initial electrostatic profiles as output. As a loss function to train the deep network, we use mean squared error between generated and simulated electrostatic potential profiles [6]. To use the electrostatic profiles generated by the neural network as the initial solution of the device simulator, additional quantities such as the carrier densities are required. Fixed-potential simulations provide these quantities. To train the neural network structure, we use PyTorch library. For the device simulation, our in-house code written in C++ is used.

In our previous work [5], we use a deep neural network which is composed of multi-layer perceptrons (MLP) as the device structure is one-dimensional. To extend it to the two-dimensional device structure, we use the convolutional neural network (CNN). In particular, we use an auto encoder whose architecture is excerpted from the deep convolutional generative adversarial network structure (DCGAN) [7]. Fig. 2 shows the neural network structure employed in this work. The trained convolutional network can be applied to not only a single structure but also a set of several devices in a restricted parameter range, as demonstrated in Section III-B.

III. Numerical results for bipolar junction transistors

We assume silicon devices at room temperature and consider both types of carriers (electrons and holes) in all of our simulations. In this preliminary study, constant carrier mobility

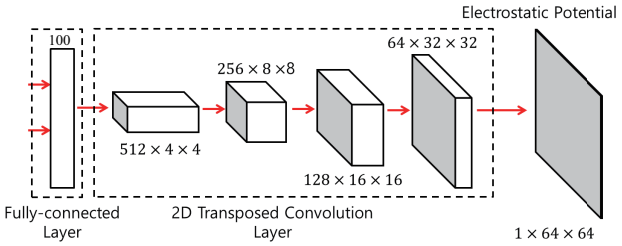


Fig. 2. Layer structure of the CNN adopted in the two-dimensional problem. The output layer generates a 64-by-64 matrix corresponding to the two-dimensional simulation domain. It is a modified version of the generator in the DCGAN in [6].

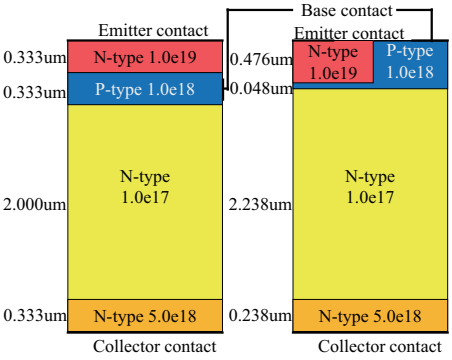


Fig. 3. Structures under consideration. A simple BJT (Left) and a more realistic BJT (Right). Numbers inside the device structures represent doping densities in cm^{-3}.

is adopted. But, more advanced mobility models can be readily used. No generation-recombination process is introduced.

A. *Fixed device structures*

Before applying our method to a set of several devices, two fixed device structures are tested in order to see the feasibility. The two BJT structures that we consider is shown in Fig. 3.

1) A simple BJT: First, we try our method on a simple BJT structure. It is a quasi-1D structure. The training data set contains 10,000 instances of 64-by-64 electrostatic potential profiles. Each instance is labeled with a pair of (V_B, V_C). The base voltage (V_B) varies from 0.0V to 0.5V. The collector voltage (V_C) is ranged from 0.0V to 0.4V. The training and validation errors are measured as functions of learning epoch as shown in Fig. 4.

After training the network with 100 epochs, we obtain a trained network to generate an approximate potential profile in the inference phase. Fig. 5 shows an example of the electrostatic potential profile generated by the trained neural network with its error. The maximum error of all test instances (not included in the training set) is lower than 21 mV. It implies that the generated potential profiles by the convolutional neural networks are close enough to the simulated profiles.

Fig. 6 shows the convergence behavior of the entire solution procedure to get the solution at the target bias condition. The target bias condition is V_B = 0.5V and V_C = 0.4V. When

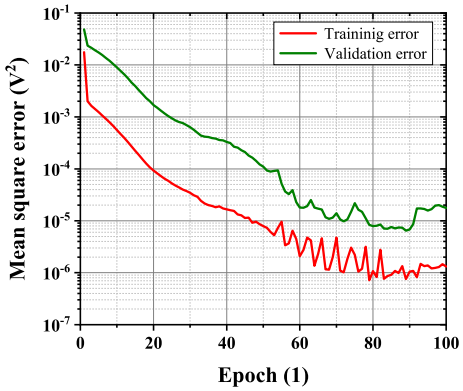
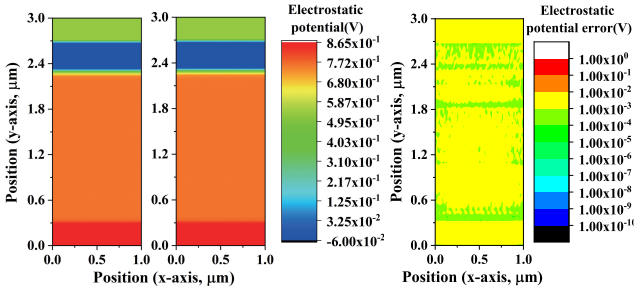


Fig. 4. Training and validation errors of a convolutional neural network, which is trained for the simple BJT structure.



(a) Numerical solution (Left) and a generated potential profile(Right). (b) Error

Fig. 5. (a) Generated potential profile and (b) its error of the simple BJT when V_B = 0.4192 V and V_C = 0.2586 V.

the maximum potential update is smaller than 10^{-10} V, the convergence criterion is satisfied. Starting with the generated potential profiles, the converged solution at the target bias condition is readily obtained in four Newton iterations with no bias ramping process. On the other hand, a conventional bias ramping with a uniform bias step ("Scratch condition" in the legend of Fig. 6) requires a larger number of Newton iterations. The number of Newton iterations can be reduced significantly with the generated potential profiles by comparing green curves and a red curve.

2) A more realistic BJT: Second, we try our method on a more realistic BJT structure. This structure has a base contact on the top surface. We configure the training data set for this structure following the same configuration of that of the first structure, and the same ranged input parameters to the first structure. The training and validation errors are shown in Fig. 7.

Fig. 8 shows an example of the generated electrostatic potential profile and its error. The maximum error of all test cases (not included in the training set) is lower than 38 mV, which is in a similar level with the simple BJT structure. It clearly demonstrates that two-dimensional doping profile does not introduce additional difficulties in the training phase. Also, its convergence behavior is shown in Fig. 9. The target bias

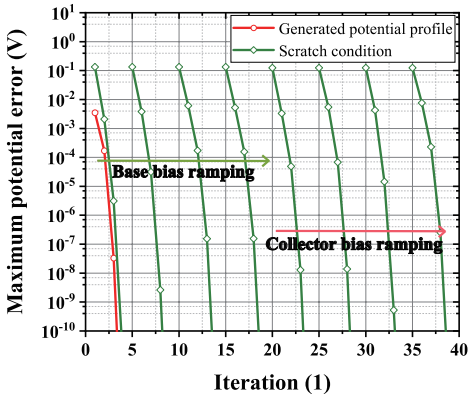


Fig. 6. Comparison of convergence behavior. The simulation with the generated initial profiles takes only four iterations for the converged solution. The target bias condition is V_B = 0.5 V and V_C = 0.4 V.

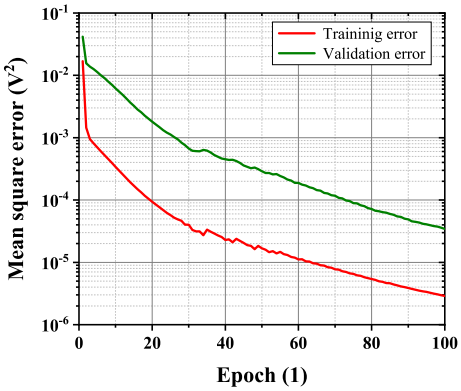
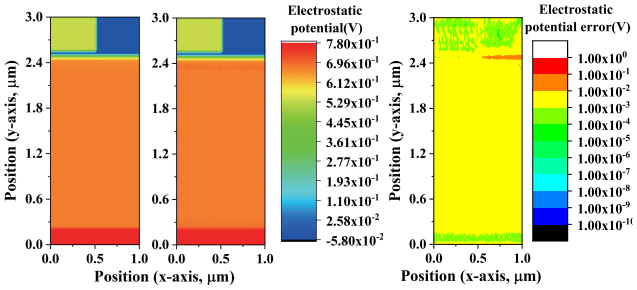


Fig. 7. Training and validation errors of a convolutional neural network, which is trained for the realistic BJT structure.



(a) Numerical solution (Left) and a generated potential profile(Right). (b) Error

Fig. 8. (a) Generated potential profile and (b) its error of the more realistic BJT when V_B = 0.4192 V and V_C = 0.2586 V.

condition is V_B= 0.5 V and V_C = 0.4 V. Again, the number of iterations can be reduced significantly by comparing green curves and a red curve.

B. Various devices

We additionally present more results on various devices. Fig. 10 shows the template of device structures considered. Variable device parameters are as follow: X_1 is the length of the emitter contact and Y_2 is the base thickness. In addition to these two geometrical parameters, the bias voltages (V_B and V_C) and the doping densities (in the base region, the collector region, and the emitter region) are also varied. These variable parameters are used as input parameters of the convolutional neural network.

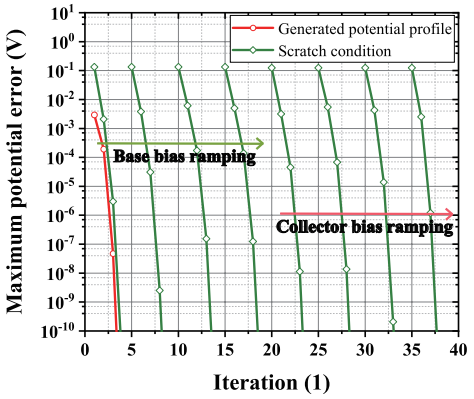


Fig. 9. Comparison of convergence behavior. The target bias condition is V_B= 0.5 V and V_C = 0.4 V.

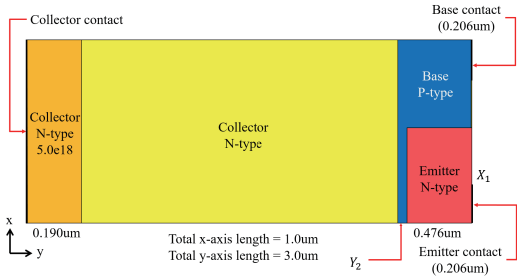


Fig. 10. Device structure under consideration. X_1 is the length of the emitter contact and Y_2 is the base thickness.

TABLE I
RANGES OF INPUT PARAMETERS

Parameter	Range
Collector voltage	$0.0\text{V} \sim 0.4\text{V}$
Base voltage	$0.0\text{V} \sim 0.5\text{V}$
X1 (Emitter length)	$0.254\mu\text{m} \sim 0.762\mu\text{m}$
Y2 (Base thickness)	$0.048\mu\text{m} \sim 0.019\mu\text{m}$
Collector doping density	$0.7 \times 10^{17}\text{cm}^{-3} \sim 1.3 \times 10^{17}\text{cm}^{-3}$
Base doping density	$0.7 \times 10^{18}\text{cm}^{-3} \sim 1.3 \times 10^{18}\text{cm}^{-3}$
Emitter doping density	$0.7 \times 10^{19}\text{cm}^{-3} \sim 1.3 \times 10^{19}\text{cm}^{-3}$

Each input parameter in the training set is randomly picked within its range. The training data set consists of 25,000 instances of these input parameters and their corresponding potential profiles. Table I shows the ranges of input parameters.

The training and validation errors are shown in Fig. 11. Errors are successfully reduced after 100 epochs. Figs. 12 and 13 show two examples of the generated electrostatic potential profiles and their errors. These examples demonstrate that the convolutional neural network can generate the potential profile

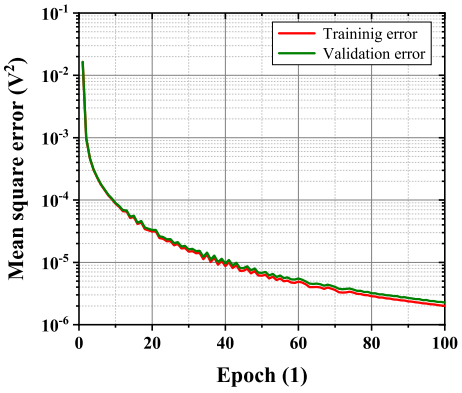


Fig. 11. Training and validation errors of a convolutional neural network which is trained for a set of various BJT structures.

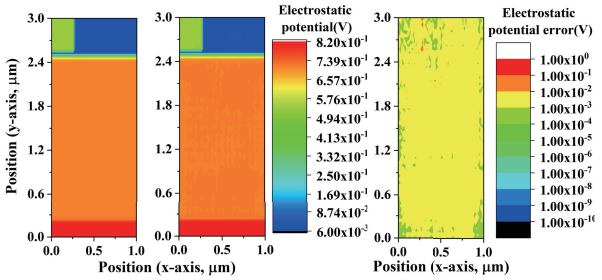


Fig. 12. Numerical solution (Left), a generated potential profile by the convolutional neural network (Center), and its error (Right). V_B = 0.5 V, V_C = 0.3 V, X_1 = 0.254 μm, Y_2 = 0.0476 μm, (Base doping)= $1.0\times10^{18}\mathrm{cm}^{-3}$, (Emitter doping)= $1.1 \times 10^{19}\mathrm{cm}^{-3}$, and (Collector doping)= $8.0 \times 10^{16}\mathrm{cm}^{-3}$.

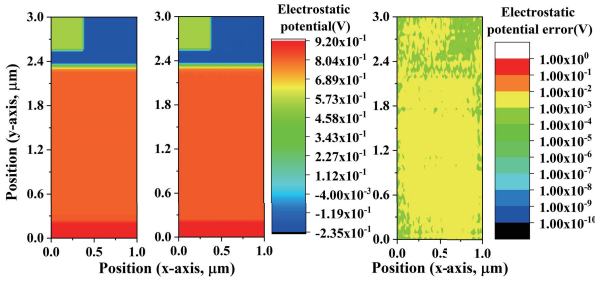


Fig. 13. Numerical solution (Left), a generated potential profile by the convolutional neural network (Center), and its error (Right). V_B = 0.25 V, V_C = 0.4 V, X_1 = 0.365 μm, Y_2 = 0.190 μm, (Base doping)= $1.3\times10^{18}\mathrm{cm}^{-3}$, (Emitter doping)= $1.0 \times 10^{19}\mathrm{cm}^{-3}$, and (Collector doping)= $1.3 \times 10^{17}\mathrm{cm}^{-3}$.

TABLE II
COMPARISON OF THE NUMBER OF NEWTON ITERATIONS

Training structure	Maximum Newton iterations	
	Generated profiles	Conventional profiles
Simple BJT	4	38
More realistic BJT	4	39
Non-fixed BJT	5	36

even when the training set contains various structures with the length variation.

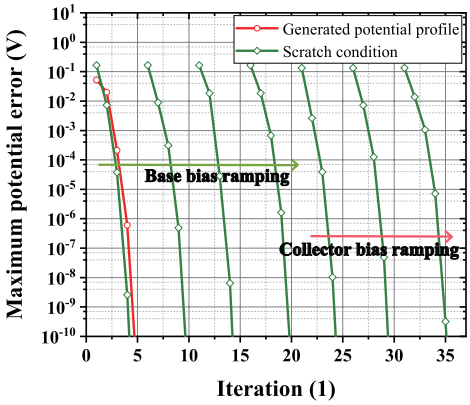


Fig. 14. Comparison of convergence behavior. The target bias condition is V_B = 0.5 V and V_C = 0.4 V. X_1 = 0.333μm, Y_2 = 0.190μm, (Base doping)= $7.7 \times 10^{17}\mathrm{cm}^{-3}$, (Emitter doping)= $1.0 \times 10^{19}\mathrm{cm}^{-3}$, and (Collector doping)= $1.2 \times 10^{17}\mathrm{cm}^{-3}$.

Fig. 14 shows the convergence behavior for an arbitrarily chosen device. Table II compares the maximum number of Newton iterations for convergence between the generated potential profiles and the conventional initial profiles, shown in Figs. 6, 9, and 14.

IV. Conclusion

We show that the trained convolutional neural networks can generate the electrostatic potential profiles which are close to the solutions. These profiles can be used as good initial solutions to reduce the number of Newton iterations. It leads to significant computational gain in a semiconductor device simulation. Extensions to more general cases will be an interesting future research topic.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2019R1A2C1086656). This work was also supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2019-0-01351, Development of Ultra Low-Power Mobile Deep Learning Semiconductor With Compression/Decompression of Activation/Kernel Data).

References

[1] W. Haensch, "Analog computing for deep learning: Algorithms, materials & architectures," in *IEDM*, 2018.
[2] J. Welser, J. Pitera, and C. Goldberg, "Future computing hardware for ai," in *IEDM*, 2018.
[3] B. Kim and M. Shin, "Machine-learning-based device optimization with tcad," in *KCS*, 2020.
[4] R. Orihara, R. Naraski, Y. Yoshinaga, Y. Morioka, and Y. Kokojima, "Approximation of time-consuming simulation based on generative adversarial network," in *ISCSA*, 2018.
[5] S.-C. Han and S.-M. Hong, "Deep neural network for generation of the initial electrostatic potential profile," in *SISPAD*, 2019.
[6] S. J. Russell and P. Norvig, *Artificial Intelligence-A Modern Approach*. Pearson Education London, 2010.
[7] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, arXiv preprint, arXiv:1511.06434.