Device-Simulation-Based Machine Learning Technique for the Characteristic of Line Tunnel Field-Effect Transistors

CHANDNI AKBAR[1,2,3], YIMING LI[1,2,3,4,5,6], (Member, IEEE), AND NARASIMHULU THOTI[1,4], (Graduate Student Member, IEEE)

[1]Parallel and Scientific Computing Laboratory, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
[2]Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
[3]Department of Electrical Engineering and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
[4]Electrical Engineering and Computer Science International Graduate Program, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
[5]Institute of Biomedical Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
[6]Center for mmWave Smart Radar System and Technologies, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan

Corresponding author: Yiming Li (ymli@nycu.edu.tw)

ABSTRACT With the rapid growth of the semiconductor manufacturing industry, it has been evident that device simulation has been considered a sluggish process. Therefore, due to downscaling of semiconductor devices, it is significantly expensive to obtain the inevitable device simulation data because it requires complex analysis of various factors. To develop a competent technique to analyze the performance of the line tunnel field-effect transistors (TFETs), the 3-D stochastic device simulation is integrated with a machine learning (ML) algorithm, named random forest regressor (RFR). Despite producing tremendous researches by the RFR model in the field of computer vision, the adoption of these ML algorithms in the field of the semiconductor industry has a lot of margin for progress. The ML-based RFR model is exploited to predict the effect of variability sources of line TFET under different biasing conditions. Results are promising and reducing the computational cost of device simulation by 99%. The prediction of effect of source variation is less than 1% as compared to the device simulation of line TFET. The application of the RFR on the line TFET device exhibits the power and flexibility of this approach because its evaluation with different bias conditions shows outstanding results.

INDEX TERMS Artificial intelligence, random forest regressor, intelligent manufacturing, machine learning, line tunnel field-effect transistors.

I. INTRODUCTION

The electrical characteristics of tunnel field-effect transistors (TFETs) outperform as compared to complementary metal-oxide semiconductors [1]. However, for commercial device manufacturing, TFETs still need to improve some specific electrical characteristics, i.e., on-state current (I_{ON}) and steep or average subthreshold swing (SS_{avg}) while maintaining controlled off-state current (I_{OFF}). Nevertheless, TFET exhibits improved I_{ON} and SS_{avg} that can inaugurate various TFET applications regarding different device structures

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

and material options [2]. The tunneling effects can be enhanced through crucial factors such as material [3], oxide (high-κ) [4], gate engineering (gate-all-around) [5], geometrical options (nanowire or nanosheets through confined width) [6], etc., [7]. In addition, the recent outbreak reveals that the utilization of ferroelectric material in TFET devices introduces the improvement of I_{ON} through internal voltage amplification [8]. There is another way to improve the tunneling probability is by producing a stronger electrical field by experiencing new and different options, such as vertical or line tunneling mechanisms [9]. Other techniques such as the utilization of 2D materials and multi-channel concepts can enhance the performance of TFET [10], [11]. Owing to these

options, we provide the promising structure of the line TFET in our recent demonstration [12]. However, there is a few prior research that can examine TFET performance exclusively by implementing emerging machine learning (ML) technology [13], [14]. For the first time, we analyze the effect of the line TFET characteristics by varying several device parameters such as source overlap length (L_{ov}), epitaxy (n) thickness (t_n), oxide thickness (t_{ox}) and work function (WK). Thus, it is vital to analyze the effect of these parameters of the line TFET by implementing emerging ML techniques.

Recently, ML is becoming visible in all research areas such as physics [15], mathematics [16], chemistry [17], etc. As we know that, ML has already paved its path in computer vision [18] and image processing [19]. But, ML has very few legitimate applications in the semiconductor industry. Therefore, in the field of semiconductor manufacturing, it has a broad way to explore its harmony with device simulation. In [20], the comparison of three different deep learning (DL) algorithms, i.e., artificial neural network, convolutional neural network and long short term memory, were implemented using device simulation data of gate-all-around silicon nanowire MOSFET and electrical characteristics were predicted through the work function fluctuation of random nanosized metal grains on MOSFET channel. Similarly, in [21], the ML algorithm was applied to predict the I_D-V_G curves obtained through device simulation of multichannel gate-all-around silicon nanosheet MOSFETs. Moreover, in [22], the DL model was implemented to predict five output features obtained from I-V/C-V curves. In [23], the DL model was investigated to predict the worst random discrete dopant configuration obtained through the device simulation. In [24], DL approach is utilized to estimate the work function fluctuation of gate-all-around silicon nanoseheet MOSFETs with a ferroelectric HZO layer. In our prior work [25], the ML-based random forest regressor (RFR) model was implemented on simulation data of the line TFET while considering only two device parameters, i.e., L_{ov} and W with the fixed biasing condition, i.e., V_D = 0.5 V and the error rate between predicted and simulated values was 5% which is recognized as a huge number in the field of ML technology.

RFR model has many advantages that make it an outstanding ML model. It emphasizes the feature selection that can help to give more importance to the valuable features and can prune the noisy/unimportant features. The other main advantage of RFR model is that it can handle non-linear data as well whereas, the other ML models have lacked this property. In this work, ML algorithms aim to overcome the computational cost and non-holistic optimization of a complex structure of the line TFET and its electrical characteristics. Instead of the derivation of conventional complex equations, ML algorithms are subsequently optimizing the possible solution for the device parameters without any specific knowledge of the device physics. In order to overcome the primary issues [26], three main contributions of this work are: (1) overcoming the complexity and ambiguity of device simulation at sub-3-nm technology node. (Nevertheless, the
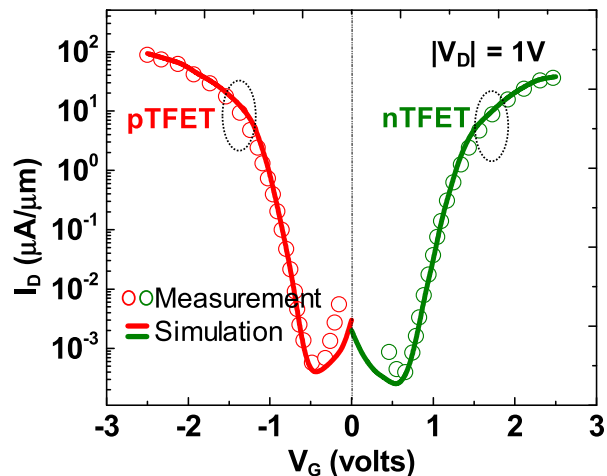

FIGURE 1. An illustration and validation of device simulations with experimentally measured data corresponding to $Si_{0.85}Ge_{0.15}$ TFET [30].

quantum confinement is crucial for the device dimensions below the effective width ($\sim$5-7nm) [31]. Our specifications that fall under the sub-3-nm technology node are L_g = 15nm, effective width >14 nm, and so on), (2) complex modeling and optimal solution without compromising the ML model's prediction accuracy, and (3) holistic and optimized solution for flexible design of the line TFET applications. In addition, our explored ML model is also able to predict the region of operation of the line TFET device by considering the I_D-V_G values. Instead of encountering few device parameters, the RFR model deals with the predictive inference using four crucial device parameters of the line TFET, i.e., L_{ov}, t_n, t_{ox} and WK.

This paper is structured as follows. In Section II, device simulation and data collection procedures are explained. Section III demonstrates the modeling of the ML model based on the device simulation. Section IV presents the results and the discussion and Section V defines the conclusions and suggests the future work.

II. DEVICE SIMULATION AND DATA GENERATION
To provide the best accuracy of device simulation, the calibration [31] is performed with experimental data, as shown in Fig. 1. It can be seen that though the calibration is performed with respect to si device the SiGe has minor variability especially in terms of tunneling mass and energy bandgap (which are crucial for tunneling probability) [11]. Nevertheless, the figure is updated by calibrating with SiGe device [31]. In this paper, our proposed scaled line TFET (SLTFET) with nanosheet geometry is utilized, as shown in Fig. 2(a). 3-D device simulations [11], [12], [28], [29] by considering the band-to-band tunneling (BTBT) model of dynamic nonlocal and trap-assisted tunneling (TAT) for effective estimation of I_{OFF} are utilized. The tunneling transport is determined by the evaluation of the BTBT model and TAT models, especially in TFETs. Here, TAT model estimates the influence of estimates the influence of trap-assisted-tunneling, which
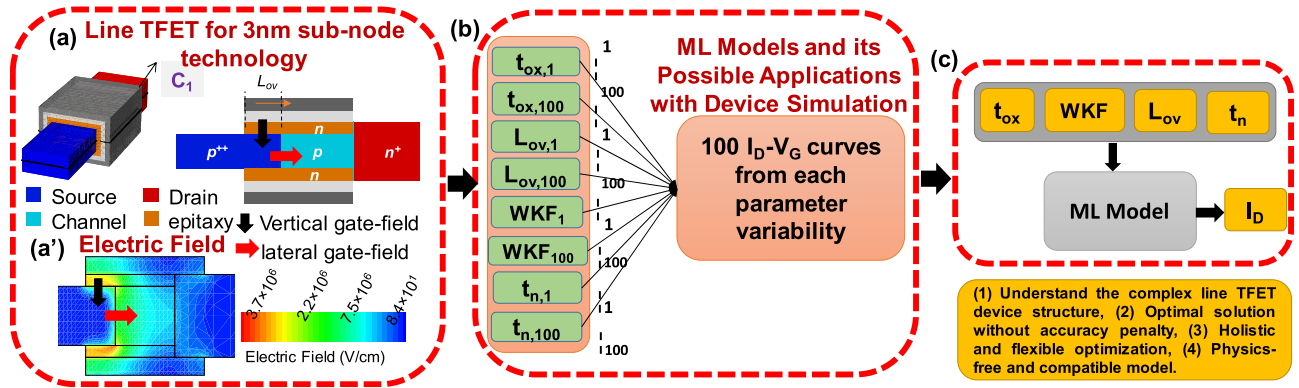
FIGURE 2. (a) An illustration of scaled line TFET for sub-3-nm technology node and (a') shows the vertical as well as lateral gate field of line TFET, (b) Various random sources and device parameter variability and 100 I_D-V_G curves obtained from each variability source, (c) A general ML model that is fed by five crucial electrical and device parameters as input and drain current is taken as a target. It also demonstrates the possible application of concatenation of the ML algorithm with the line TFET device simulation.

is major contribution for the off-state current that arises during zero-gate bias. The tunneling parameters according to the material options ($Si_{1-x}Ge_x$ with x as Ge fraction of 0.4) are calibrated and included in the device simulations. i.e. effective mass based on conduction and valence band density of states (m_c and m_v) are $0.328m_0$ and $0.421m_0$ for $Si_{0.6}Ge_{0.4}$. In addition, the Kane's model (BTBT) direct and indirect path parameters Adir, Bdir, Aindir, and Bindir for $Si_{0.6}Ge_{0.4}$ are $1.341 \times 10^{20} cm^{-3}S^{-1}$, 54.05 $MVcm^{-1}$, 2.44×10^{15} $cm^{-3}S^{-1}$, and 16.8 $MVcm^{-1}$, respectively. More discussion on these calibrations can be found in our recent articles [11], [25].

Fig. 2 demonstrates the design of SLTFET using an n-epitaxial layer over the channel and source and to improve the vertical gate-field via a L_{ov}. The working principle of the explored TFET depends on both vertical (source-epitaxy) and lateral (source-channel) tunneling mechanisms. During the off-state; i.e. drain voltage (V_D) = 0.5 V and gate voltage (V_G) = 0 V, the tunneling length (λ) is longer to have reasonable BTBT across both the junctions. As long as the applied potential increases (on-state; $V_D = V_G = 0.5$ V) the BTBT rate exponentially increases for the generated gate vertical- and lateral-fields as shown in Fig. 2(a'). Here, the magnitude of the vertical field is stronger than the lateral field as long as L_{ov} exists. It is to be noticed that the key parameters of TFET depend on energy bandgap, effective mass, tunneling length, and so on, which are related to material engineering. However, the geometrical options with respect to structure selection influence L_{ov}, t_n, oxide thickness (t_{ox}), work function, etc. [30]. Hence, we investigate the performance of TFET through geometrical options rather than material considerations. The significance of each device parameters are described below.

A. SIGNIFIANCE OF OVERLAPPING LENGTH (L_{OV})

The significance of L_{ov} is to modulate the vertical gate-field as well as vertical tunneling via p^{++}-n (refer to Fig. 1(a)). The factor of L_{ov} helps to improve the area of tunneling (A_{tun}),

TABLE 1. List of the line TFET device parameters and their ranges with their respective step size explored in ML RFR model.

Device Parameters	Range
L_{ov} (nm)	0.1 to 10 : Step size = 0.1 nm
WK (eV)	4.2 to 4.5: Step size = 0.025 eV
t_n (nm)	1 to 5 : Step size = 0.025 nm
t_{ox} (nm)	1 to 6 : Step size = 0.05 nm

i.e. $A_{tun} = L_{ov} * 2(W + t_n)$, where t_n and W are the thickness and width of the channel. This refers to the proportionality of tunneling or A_{tun} with respect to L_{ov} and device dimensions such as t_n and W. Here t_n is varying as well as W is varied accordingly, as listed in Table 1.

B. SIGNIFIANCE OF EPITAXIAL THICKNESS (T_N)

The value of t_n is also significant because the band alignment between p^{++}-n will determine the tunneling barrier length (λ). An appropriate band alignment is responsible for a greater tunneling rate [29]. Other material factors that will influence the tunneling rate are t_{ox} and WK.

C. SIGNIFIANCE OF OXIDE THICKNESS (T_{OX})

The factor of t_{ox} highly influences λ, as $\lambda = \sqrt{(\varepsilon_{ns} t_{ox} t_{ns})/\varepsilon_{ox}}$, where ε_{ns} and ε_{ox} are the permittivity of nanosheet and gate-oxide, respectively.

D. SIGNIFIANCE OF WORK FUNCTION (WK)

The WK would make an effect on the subthreshold operation of the device that influences with a low threshold (V_t) and deviation in SS values. Here, titanium nitride (TiN) is used for making WK as an n-type device structure. The work function range of 4.2-4.4eV is considered to maintain a high I_{ON}/I_{OFF} ratio. This is because the TFETs suffer with low on-current

**TABLE 2.** List of the hyperparameters utilized to train the ML based RFR model.

Hyperparameters	Value
Number of Trees (*N Estimators*)	50
Maximum Depth	None
Criterion	MSE
Random State	42

FIGURE 3. An illustration of the ML algorithm explored for the line TFET device simulation data. (a) The RFR algorithm based on 50 trees with varying depth of each tree to predict the I_D curve. (b) Data is splitting between the training and the testing set. Input and output/target parameters are also defined.

at high WK. Therefore, it is meaningful to consider a low work function range that implies steeper band banding ($\Delta\phi$), proportionally high tunneling rate or on-current. In reality, it has been identified that the variation in WK from low to high with meaningful offset can be achieved through plasma-ion implantation [33].

Hence we have specifically input these parameters into the ML RFR model to understand the device structure of line TFET. Furthermore, to improve the tunneling rate the explored SLTFET is utilized with hetero-structure having $Si_{0.6}Ge_{0.4}$ as the source and the rest with that of Si. The data obtained through the device simulation is fed into the RFR model such as the input features are V_G, L_{ov}, t_n, t_{ox} and *WK*. These input features are composed of different ranges and each device parameter is generating 100 $I_D - V_G$ curves, as shown in Fig. 2(b). Therefore, our explored RFR model is intrigued using 400 I_D-V_G curves. Similarly, the output feature is I_D. After specifying the input and the output features for the ML model, it is necessary to split the data into the training and the testing sets. The split is user-specific as well as ML model-dependent. Furthermore, the normalization of a dataset is performed to compose the whole data into the standardized range to improve the accuracy of the ML model. Fig. 2(c) illustrates the input and output from the ML model as well as the possible ML application for the line TFET simulated data.

III. MODELING OF MACHINE LEARNING ALGORITHMS

Since 40 years ago, researchers have been struggling to formulate the simple equation of the complex structure of semiconductor devices. Therefore, to make a general model based on multiple hyperparameters, the RFR algorithm is implemented that can work for the SLTFET with the given input and output vectors. The ML-based RFR model is tuned with the help of various hyperparameters as listed in Table 2. RFR model having a bunch of parallel decision trees has the advantage to make a flexible model by varying the hyperparameters. In this work, the RFR model is based on 50 decision trees, as shown in Fig. 3(a). Before feeding the training set into the RFR model, the training dataset is preprocessed such as shuffling, normalization, and splitting of data into an
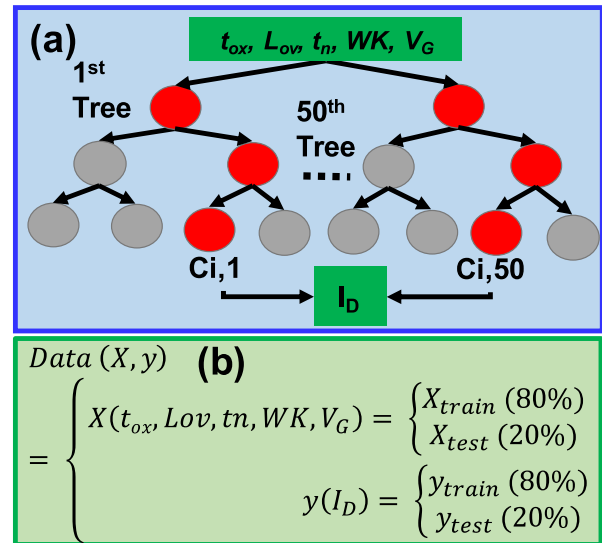
appropriate ratio for training and testing the model. In general, the testing data remain unknown to the ML model. Firstly, the $I_D - V_G$ curves are shuffled so that the ML model is able to be trained from each possible curve from all ranges of device parameters. Secondly, the data is normalized to eliminate the outliers. The linear normalization is performed by subtracting the data from its mean value and then dividing it by its standard deviation. After normalization, all the training dataset is in the range of -1 to 1. Thirdly, the data is split into 80% for the training set and 20% for the testing set. The partition of data into the training and testing set is presented in Fig. 3(b). Notably, while training the RFR model and during the evaluation of the trained RFR model, the mean squared error (MSE) value is calculated as a loss function. Moreover, in this work, the R^2-score is also considered as the source of evaluation of the trained ML model. The higher value of the R^2-score shows that the input variables are perfectly correlated, whereas, a value closer to 0 shows that the ML model is not valid and suffering from many problems related to train/test data split, noise in the data, unavailability of tuned hyperparameters of the ML model, and so on. Our approach is to use the RFR model to predict the I_D curves from the given device parameters and the electrical characteristics. For this purpose, the RFR model is trained with four different datasets such as (1) the 400 $I_D - V_G$ curves are collected by varying one device parameter at a time while fixing the rest of the three parameters, (2) among four crucial device parameters, 200 $I_D - V_G$ curves are generated by varying L_{ov} and t_n with each other to show the correlation between these two parameters, (3) the electrical characteristics are investigated by varying the device parameters as well as drain voltage (V_D), and (4) approximately 2500 fluctuated devices

TABLE 3. Comparison of various data split and their corresponding accuracy (R²-score) of ML-RFR model.

Case	Data Split Percentage		Training R^2-Score (%)
	Train	Test	
a	50	50	98.7
b	60	40	99.2
c	70	30	99.7
d	80	20	99.9
e	90	10	99.9

FIGURE 4. An illustration of performance of the ML-RFR model using RMSE value by varying the number of trees. It shows that the minimum RMSE value is achieved at 50 number of trees.

are generated to collect the $I_D - V_G$ curves with the variation of all the device parameters to exhibit the relationship among WK, t_{ox}, t_n and L_{ov}.

This proposed idea is more efficient and accurate because many input features are investigated with respect to the hidden parameters of the ML model and the output reflects the collective as well as individual effect of all device parameters, i.e., WK, V_D, t_{ox}, t_n and L_{ov}. Notably, Table 1 lists the varying range of the device parameters of the SLTFET device utilized in the modeling of the RFR algorithm.

While optimizing the RFR model, there is no certainty to obtain the controllable range of real semiconductor device parameters because the training of the model does not relate to any device physics and may try to find the optimal solution for all the input features. To obtain the well-regulated range of the predicted output, it is necessary to scale the input features using Python's library, i.e., Scikit-learn [18]. All the experiments are operated in Python console on the computer with Intel i7-10700K CPU (3.97 GHz) and 32.0 GByte RAM.

IV. RESULTS AND DISCUSSION

50 trees-based RFR model using five input nodes and having an adjustable depth for each decision tree, has been implemented via Python's library, i.e., Scikit-learn. While implementing the RFR model, the number of trees and the depth of each tree has been determined by the number of $I_D - V_G$ curves to learn the relationship between the device parameters and the target value. In addition, to achieve robust results, the number of training samples and hyperparameters must be in an appropriate manner to avoid both overfitting and underfitting of the RFR model.

Generally, the dataset is split into three subsets, i.e., training set, testing set and validation set. Validation and test set play an important role but sometimes validation set is not required when dataset is small and the error rate for training as well as testing data are in good agreement. Before feeding the dataset into the ML model, the dataset is split into the training and testing set. It is a complicated task especially when the dataset is small. The training and testing data split affects the various attributes of the ML model. For example, the accuracy and the best fitting of the ML model depend on the appropriate selection of train/test split. The appropriate train/test split tune the hyperparameters in such
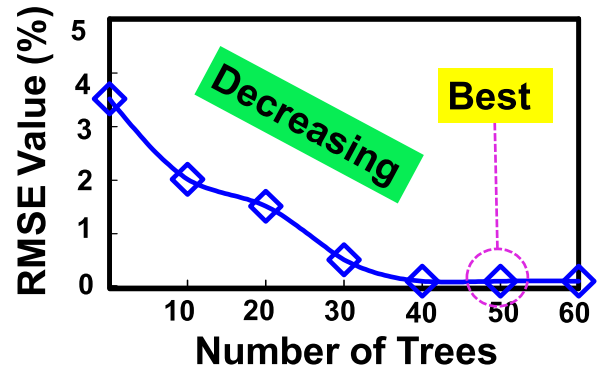
a way so that it can produce the best accurate predictive ML model. Therefore, while investigating the 400 I_D-V_G curves, to determine the best splitting ratio, the R²-score of the training of the explored ML model is calculated by considering varying dataset splits, as listed in Table 3. It can be seen that the most appropriate split for our small dataset is the case (d) and case (e). Therefore, randomly 320 I_D-V_G curves are selected as a training dataset and the rest of the curves are utilized for the evaluation of the trained RFR model. The RFR model is trained by considering the case *d*. The tunned hyperparameters utilized by case *d* are listed in Table 2. Moreover, Fig. 4 shows that 50 trees have the minimum root mean squared error (RMSE) for this dataset. Thus, our explored RFR model is constructed by using 50 number of trees. During the training of the RFR model, we stop splitting the nodes in the trees based on the generalization of the performance of the MSE value. For example, if the MSE value remains the same for the previous nodes, then the output is taken from that node. We repeat this process for several decision trees and output is obtained by taking an average of all the acceptable outcomes. As it has been known that the RFR model is stochastic, therefore, its performance can vary by the initial random parameter values. To avoid overfitting and for the best accuracy performance, the RFR model is trained and evaluated by initializing with different parameter values. The training of the ML model is physics-free, i.e., it is working without any knowledge of the device physics. I_D-V_G curve is evaluated by using the RMSE value and R²-score. RMSE value measures the difference between the true value/simulated value and the predicted value (output from the RFR model). The lesser the RMSE value, the more accurate is the performance of our explored model. Similarly, R²-score is a statistical measure that reflects the fitness of the ML model. Its value ranges between 0 to 1. Moreover, a value closer to 1 exhibit the best performance of the ML model and vice versa. Notably, the training and the testing of the RFR model using different device parameters are illustrated in Fig. 5. In this proposed idea, each device parameter has different ranges, however, in Figs. 5(a)-(h), solid line (-) shows the
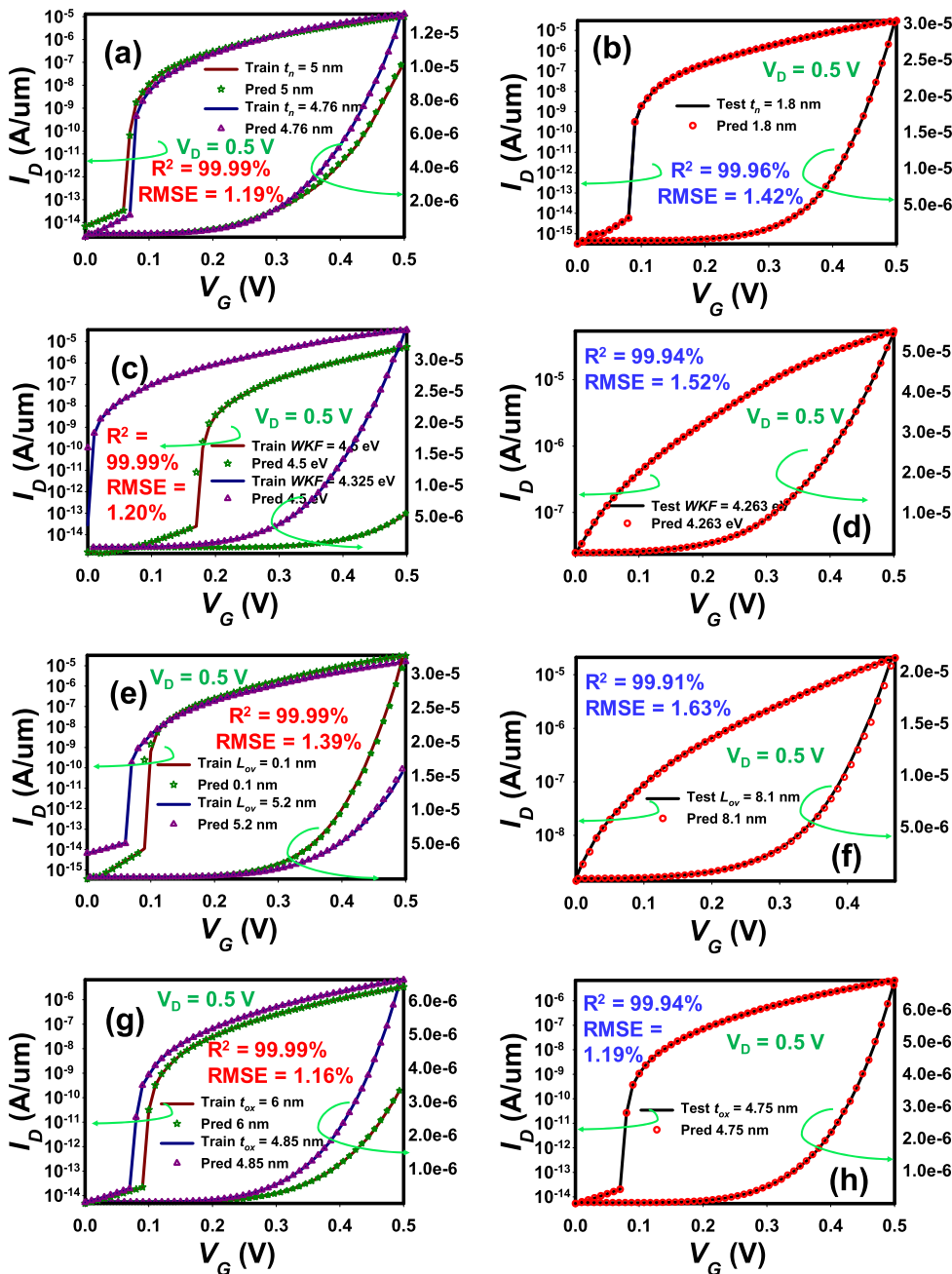
FIGURE 5. An illustration of training and testing of the RFR model. (a) and (b) represent the I_D-V_G curves during training and testing of the line TFET by considering device parameter, i.e., t_n = 1 to 5 nm with 0.1 step size, obtained through device simulation, respectively, (c) and (d) represent the I_D-V_G curves during training and testing of the line TFET by considering device parameter, i.e., WK = 4.2 to 4.4 eV with 0.025 step size, obtained through the device simulation, respectively, (e) and (f) represent the I_D-V_G curves during training and testing of the line TFET by considering device parameter, i.e., L_{ov} = 0.1 to 10 nm with step size of 0.1 nm, obtained through device simulation, respectively, and (g) and (h) represent the I_D-V_G curves during training and testing of the line TFET by considering device parameter, i.e., t_{ox} = 1 to 6 nm with 0.05 step size, obtained through the device simulation, respectively.

simulated data and marker (o) represents the predicted value from our explored RFR model. It can be noted that the I_D-V_G curve fitting outperforms for all the device parameters and due to the rigorous R^2-score, the RMSE value is diminished as well. Therefore, it can be concluded from Fig. 5 that, our

explored RFR model learned the complex equations of the SLTFET for the given range of device parameters. Moreover, our well-trained RFR model can predict the I_D-V_G curves for the unknown device parameters but in the same perturbation range of simulated data.
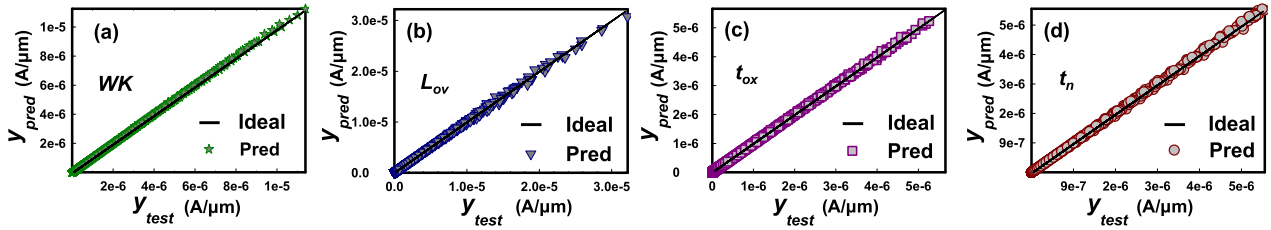
FIGURE 6. Linear regresssion plots of the testing dataset obtained through the evaluation of the RFR model using all explored device parameters. (a), (b), (c) and (d) represent the comparison of the predicted and the simulated values of WK, L_{ov}, t_{ox} and t_n, respectively.
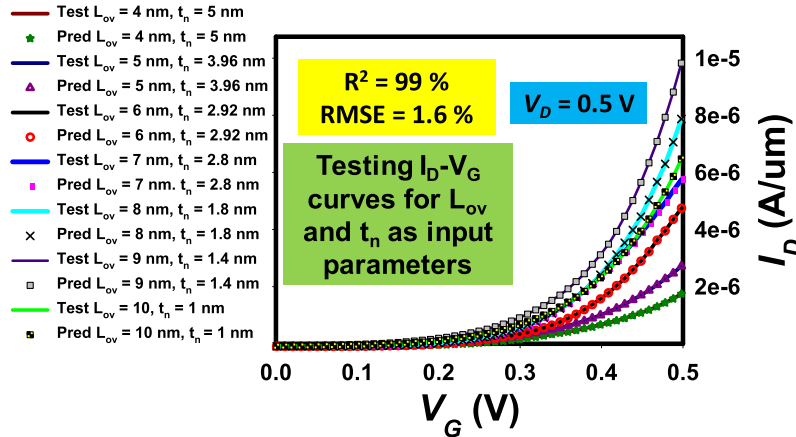

FIGURE 7. An illustration of the combined effect varying ranges of L_{ov} and t_n. It also shows the effect of these two device parameters on the prediction of I_D-V_G curves. It shows that the testing of the combined effect is two device parameters also outperform in terms of RMSE value and R^2-score for fixed biasing condition, i.e., $V_D = 0.5$ V.

Fig. 6 illustrates the output regression lines of the RFR model. It can be seen that Figs. 6(a)-(d) exhibit the linear performance for all explored device parameters by comparing the simulated values and the predicted values. Moreover, in Fig. 6, the relationship between the predicted and device simulated I_D values is approximately linear which shows that the predicted I_D values are close enough to the simulated (test) values. Therefore, it can be seen that our explored RFR model outperforms in terms of accuracy.

After exhibiting the effect of source of variation independently using 400 fluctuated devices, the second dataset is investigated, i.e., 200 fluctuated devices by varying L_{ov} and t_n simultaneously (WK, t_{ox} remain constant) to study the relationship between them. Fig. 7 presents the I_D-V_G curves obtained through the device simulation as well as the prediction via the ML-RFR model. It can be observed that the R^2-score is approximately equal to 99% and the RMSE value is very close to 1% as well. Thus, it can be concluded that our explored ML-RFR model performs well with two device parameters as well.

Thirdly, in order to demonstrate the compatibility and physics-free modeling of our explored RFR model, the model is trained and evaluated by different biasing conditions such as $V_D = 0.5, 0.05$, and 0.005 V. It can be seen from Fig. 8 that the prediction of the I_D-V_G curve is outstanding in terms of RMSE value and R^2-score. In short, it can be concluded that
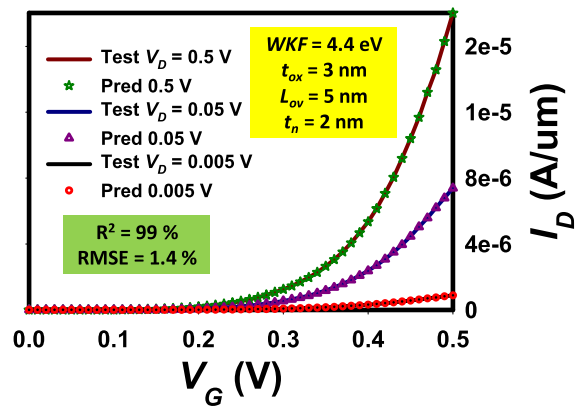

FIGURE 8. An illustration of testing of the RFR model by predicting the I_D-V_G curves of the line TFET by considering device parameters, i.e WK = 4.4 eV, $t_{ox} = 3$ nm, $L_{ov} = 5$ nm and $t_n = 2$ nm, having different biasing conditions, i.e., $V_D = 0.5, 0.05$ and 0.005 V. It shows that our well-trained model performs well on the unknown biasing conditions as well.

ML modeling is physics-free and does not require an exact equation to predict the target values.

Furthermore, to establish the relationship between all the explored device parameters, the model is trained by splitting the 2500 curves into 80% for training and 20% for testing the RFR model. Before feeding into the RFR model, the dataset goes through the preprocessing steps (as we already
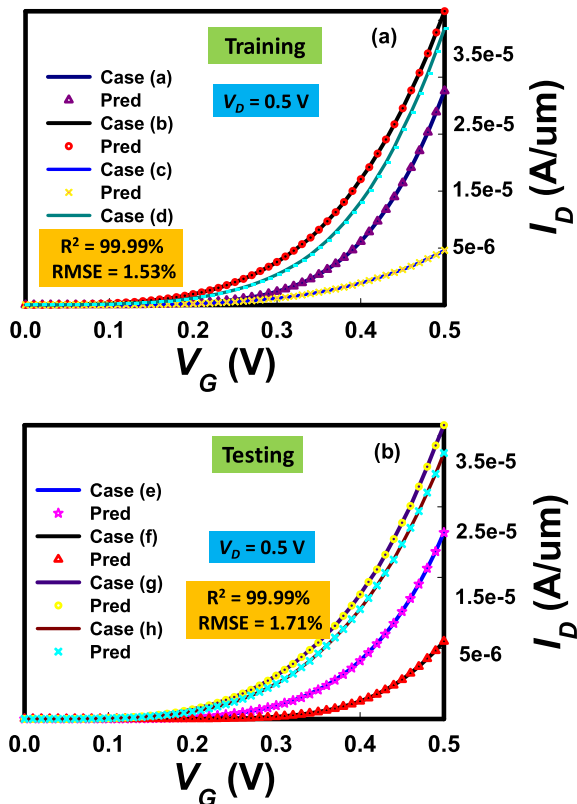
FIGURE 10. An illustration of evaluation of the well-trained ML-RFR model by testing through the random device parameters. The error rate for L_{ov} = 0.15 nm, t_{ox} = 3.04 nm, t_n = 2.04 nm, WK = 4.44 eV, is less than 1%. The difference between the ideal line and the scattered predicting points shows that it is possible to evaluate the model using any value within the specific range of device parameters.

FIGURE 9. An illustration of training and testing of the RFR model with varying all explored device parameters, i.e., L_{ov}, WK, t_n, and t_{ox}, by predicting the I_D-V_G curves of the line TFET. (a) the training of the RFR model and four I_D-V_G curves are illustrated. The detail of case (a) −case (d) is shown in Table 4. (b) shows the testing of the RFR model and for the sake of visualization four cases are shown. Detail of the device parameters of the testing curves is listed in Table 4.

TABLE 4. List of the device parameters shown (in Fig. 9) by training and testing of the RFR model.

Cases	Device Parameter
a	L_{ov} = 0.5 nm, t_{ox} = 2.5 nm, WK = 4.4 eV, t_n = 2 nm
b	L_{ov} = 2.5 nm, t_{ox} = 3 nm, WK = 4.275 eV, t_n = 2 nm
c	L_{ov} = 5 nm, t_{ox} = 4 nm, WK = 4.4 eV, t_n = 3.5 nm
d	L_{ov} = 5 nm, t_{ox} = 3 nm, WK = 4.3 eV, t_n = 1.5 nm
e	L_{ov} = 4.5 nm, t_{ox} = 3.5 nm, WK = 4.4 eV, t_n = 2 nm
f	L_{ov} = 3.5 nm, t_{ox} = 3 nm, WK = 4.5 eV, t_n = 2 nm
g	L_{ov} = 5 nm, t_{ox} = 1 nm, WK = 4.4 eV, t_n = 4 nm
h	L_{ov} = 5 nm, t_{ox} = 3 nm, WK = 4.375 eV, t_n = 3 nm

discussed). The same hyperparameters are utilized in training the RFR model except for the number of trees. In this training, 100 number of trees are explored for the converged solution. It can be observed from Fig. 9 that the training and the testing of the RFR model outperform. The RMSE value and the R^2-score show that the performance of the RFR model by varying all the device parameters is approximately similar to Fig. 5 and Fig. 7. Therefore, we can conclude that the relationship between all the explored device parameters is uncorrelated with each other. Notably, for the sake of visualization, out of 2500 fluctuated devices, only four I_D-V_G curves are shown in Fig. 9, for the training and the testing of the ML model.

Lastly, to demonstrate the holistic and flexibility of our explored RFR model, the model is tested with the randomly generated device parameters. Firstly, after training the model with the specific range (listed in Table 1) of device parameters, the model is evaluated with unknown device parameters. Thus, we have tested our trained ML model with random values (unknown to the model) as shown in Fig. 10, L_{ov} = 0.15 nm, t_{ox} = 3.04 nm, t_n = 2.04 nm, WK = 4.44 eV. Notably, these device parameter values are not included in our simulated dataset, although, these values lie in the range

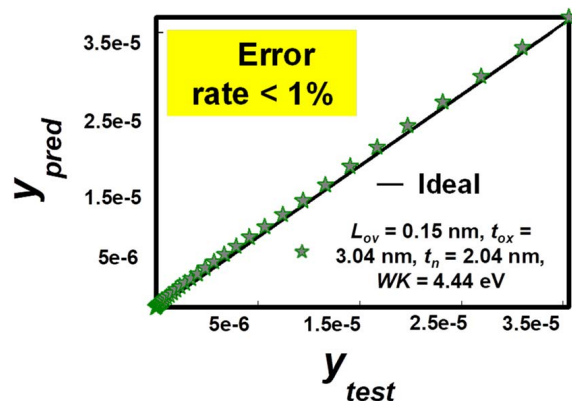of our explored parameters. The comparison is established between the simulated values and the predicted values for random cases. The regression line shows the ideal scenario and the closeness of predicted values to the ideal line exhibits the outstanding performance of the ML model. Fig. 10 concludes that the evaluation of our explored ML model using randomly selected parameters outperforms. The error rate between the predicted and the tested values are not greater than 1% which is considered a remarkable achievement in the field of ML and semiconductor device simulation. Moreover,

TABLE 5. List of simulated and predicted minimum subthreshold slope (SS$_{min}$) obtained from device simulation and ML algorithm, respectively.

Device Parameters	Simulated SS$_{MIN}$ (mV/dec)	Predicted SS$_{MIN}$ (mV/dec)
L_{ov} (1 to 10 nm)	~29	~29.14
WK (4.2 to 4.5 eV)	~32	~31.21
t_n (1 to 5 nm)	~39	~39.02
t_{ox} (1 to 6 nm)	~31	~31.05

after accurate prediction of the ML-RFR model, the crucial parameter, i.e., minimum subthreshold slop (SS_{MIN}) is extracted from the simulated data as well as from the predicted I_D-V_G curves. The extracted SS_{MIN} of line TFET device and predicted values are listed in Table 5.

Our explored RFR model takes 150 seconds to be well-trained and 20 seconds to be evaluated. Whereas, device simulation takes 5 hours to generate 100 I_D-V_G curves with one device parameter at a time. Hence, it can be concluded that ML modeling can accelerate the complex device simulation with an error rate of less than 1% and reduction of computational cost by around 99%. Moreover, innovative ML techniques can model the complex device structure and can find the optimal solution easily. Moreover, our trained model can accelerate the fabrication process because TFET is a complex and time-consuming device simulation process due to its quantum tunneling models. Therefore, predicting the I_D-V_G curve from our well-trained ML model is reliable to accelerate as well as minimize the computational cost of the fabrication process by demonstrating the electrical characteristics of any specific parameter within seconds.

V. CONCLUSION

In this work, the ML algorithm has been utilized to optimize the solution for the complex device simulation of the line TFET by training the RFR model with each possible fluctuated device having a specific range of device parameters. Five crucial device parameters, i.e., L_{ov}, t_{ox}, t_n, WK and V_D were explored to predict the I_D variation. Therefore, from the predictive results, it has been concluded that the ML algorithm is an efficient and flexible approach to predict the behavior of the line TFET for the sub-3-nm technology node. In addition, it performs well by evaluating the device parameters on other bias conditions. Therefore, it has been shown that our explored ML model is physics-free and high compatible with other device conditions as well. Furthermore, it has been accepted that the accuracy of the predictive model is far better than the human expert's optimization algorithms using device simulation. Our explored RFR model converges faster than the other traditional algorithms. The R^2-score of the well-trained and evaluation model is above 99%, similarly, the

error rate for training and testing the line TFET simulated data is less than 1% which is considered computationally efficient. Moreover, it has been concluded that all the explored device parameters are independent of each other and a generalized algorithm has been modeled for the line TFET device with specific device parameters which will be extended further in a near future by adding more material parameters as well as the process voltage temperature variations.

REFERENCES

[1] U. E. Avci, R. Rios, K. J. Kuhn, and I. A. Young, "Comparison of power and performance for the TFET and MOSFET and considerations for P-TFET," in *Proc. 11th IEEE Int. Conf. Nanotechnol.*, Aug. 2011, pp. 869–872, doi: 10.1109/NANO.2011.6144631.

[2] S. Kumar, K. Nigam, S. Chaturvedi, A. I. Khan, and A. Jain, "Performance improvement of double-gate TFET using metal strip technique," *Silicon*, vol. 14, no. 4, pp. 1759–1766, Feb. 2022, doi: 10.1007/s12633-021-00982-z.

[3] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L. Wernersson, "Vertical InAs/GaAsSb/GaSb tunneling field-effect transistor on Si with S = 48 mV/decade and I$_{on}$ = 10 μA/μm for I$_{off}$ = 1 nA/μm at V$_{ds}$ = 0.3 V," in *IEDM Tech. Dig.*, Dec. 2016, pp. 500–503, doi: 10.1109/IEDM.2016.7838450.

[4] K. Boucart and A. M. Ionescu, "Double-gate tunnel FET with high-K gate dielectric," *IEEE Trans. Electron Devices*, vol. 54, pp. 1725–1733, Jul. 2007, doi: 10.1109/TED.2007.899389.

[5] A. S. Verhulst, B. Sorée, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken, "Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor," *J. Appl. Phys.*, vol. 107, no. 2, Jan. 2010, Art. no. 024518, doi: 10.1063/1.3277044.

[6] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson, "Scaling of vertical InAs–GaSb nanowire tunneling field-effect transistors on Si," *IEEE Electron Device Lett.*, vol. 37, no. 5, pp. 549–552, May 2016, doi: 10.1109/LED.2016.2545861.

[7] X. Liu, K. Ma, Y. Wang, M. Wu, J.-H. Lee, and X. Jin, "A novel high Schottky barrier based bilateral gate and assistant gate controlled bidirectional tunnel field effect transistor," *IEEE J. Electron Devices Soc.*, vol. 8, pp. 976–980, 2020, doi: 10.1109/JEDS.2020.3020920.

[8] A. Saeidi, T. Rosca, E. Memisevic, I. Stolichnov, M. Cavalieri, L.-E. Wernersson, and A. M. Ionescu, "Nanowire tunnel FET with simultaneously reduced subthermionic subthreshold swing and off-current due to negative capacitance and voltage pinning effects," *Nano Lett.*, vol. 20, no. 5, pp. 3255–3262, May 2020, doi: 10.1021/acs.nanolett.9b05356.

[9] A. Afzalian, G. Doornbos, T. M. Shen, M. Passlack, and J. Wu, "A high-performance InAs/GaSb core-shell nanowire line-tunneling TFET: An atomistic mode-space NEGF study," *IEEE J. Electron Devices Soc.*, vol. 7, pp. 111–117, 2019, doi: 10.1109/JEDS.2018.2881335.

[10] S. Das, A. Prakash, R. Salazar, and J. Appenzeller, "Toward low-power electronics: Tunneling phenomena in transition metal dichalcogenides," *ACS Nano*, vol. 8, no. 2, pp. 1681–1689, Jan. 2014, doi: 10.1021/nn406603h.

[11] N. Thoti and Y. Li, "Influence of fringing-field on DC/AC characteristics of Si$_{1-X}$Ge$_X$ based multi-channel tunnel FETs," *IEEE Access*, vol. 8, pp. 208658–208668, 2020, doi: 10.1109/ACCESS.2020.3037929.

[12] N. Thoti and Y. Li, "P-SiGe nanosheet line tunnel field-effect transistors with ample exploitation of ferroelectric," *Jpn. J. Appl. Phys.*, vol. 60, no. 5, Apr. 2021, Art. no. 054001, doi: 10.35848/1347-4065/abf13e.

[13] CR Formatted M. Suguna, V. Charumathi, N. B. Balamurugan, M. Hemalatha, and D. S. Kumar, "Machine learning-based multi-objective optimisation of tunnel field effect transistors," *Silicon*, Apr. 2022, doi: 10.1007/s12633-022-01841-1.

[14] G. Wang, S. Wang, L. Ma, G. Wang, J. Wu, X. Duan, S. Chen, and H. Liu, "Optimization and performance prediction of tunnel field-effect transistors based on deep learning," *Adv. Mater. Technol.*, vol. 7, no. 5, May 2022, Art. no. 2100682, doi: 10.1002/admt.202100682.

[15] M. Aykol, C. B. Gopal, A. Anapolsky, P. K. Herring, B. van Vlijmen, M. D. Berliner, M. Z. Bazant, R. D. Braatz, W. C. Chueh, and B. D. Storey, "Perspective—Combining physics and machine learning to predict battery lifetime," *J. Electrochemical Soc.*, vol. 168, no. 3, Mar. 2021, Art. no. 030525, doi: 10.1149/1945-7111/abec55.

[16] A. Kumar, S. Loharch, S. Kumar, R. P. Ringe, and R. Parkesh, "Exploiting cheminformatic and machine learning to navigate the available chemical space of potential small molecule inhibitors of SARS-CoV-2," *Comput. Struct. Biotechnol. J.*, vol. 19, pp. 424–438, 2021, doi: 10.1016/j.csbj.2020.12.028.

[17] C. C. John, V. Ponnusamy, S. K. Chandrasekaran, and N. R, "A survey on mathematical, machine learning and deep learning models for COVID-19 transmission and diagnosis," *IEEE Rev. Biomed. Eng.*, vol. 15, pp. 325–340, 2022, doi: 10.1109/RBME.2021.3069213.

[18] W. Khan, K. Crockett, J. O'Shea, A. Hussain, and B. M. Khan, "Deception in the eyes of deceiver: A computer vision and machine learning based automated deception detection," *Expert Syst. Appl.*, vol. 169, May 2021, Art. no. 114341, doi: 10.1016/j.eswa.2020.114341.

[19] M. Pedram, S. A. Rokni, R. Fallahzadeh, and H. Ghasemzadeh, "A beverage intake tracking system based on machine learning algorithms, and ultrasonic and color sensors," in *Proc. 16th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, Apr. 2017, pp. 313–314, doi: 10.1145/3055031.3055065.

[20] C. Akbar, Y. Li, and W.-L. Sung, "Deep learning algorithms for the work function fluctuation of random nanosized metal grains on gate-all-around silicon nanowire MOSFETs," *IEEE Access*, vol. 9, pp. 73467–73481, 2021, doi: 10.1109/ACCESS.2021.3079981.

[21] C. Akbar, Y. Li, and W. L. Sung, "Machine learning aided device simulation of work function fluctuation for multichannel gate-all-around silicon nanosheet MOSFETs," *IEEE Trans. Electron Devices*, vol. 68, no. 11, pp. 5490–5497, Nov. 2021, doi: 10.1109/TED.2021.3084910.

[22] K. Mehta and H.-Y. Wong, "Prediction of FinFET current-voltage and capacitance-voltage curves using machine learning with autoencoder," *IEEE Electron Device Lett.*, vol. 42, no. 2, pp. 136–139, Feb. 2021, doi: 10.1109/LED.2020.3045064.

[23] J. Lee, P. Asenov, M. Aldegunde, S. M. Amoroso, A. R. Brown, and V. Moroz, "A worst-case analysis of trap-assisted tunneling leakage in DRAM using a machine learning approach," *IEEE Electron Device Lett.*, vol. 42, no. 2, pp. 156–159, Feb. 2021, doi: 10.1109/LED.2020.3046914.

[24] R. Butola, Y. Li, and S. R. Kola, "Deep learning approach to estimating work function fluctuation of gate-all-around silicon nanosheet MOSFETs with a ferroelectric HZO layer," in *Proc. EDTM*, Apr. 2022, pp. 232–234, doi: 10.1109/EDTM50988.2021.9421030.

[25] C. Akbar, N. Thoti, and Y. Li, "Machine learning approach to predicting tunnel field-effect transistors," in *Proc. Int. Symp. VLSI Technol., Syst. Appl. (VLSI-TSA)*, Apr. 2021, pp. 70–71, doi: 10.1109/VLSI-TSA51926.2021.9440136.

[26] (2020). *International Roadmap for Devices and Systems—More than Moore—IRDS*. [Online]. Available: https://irds.ieee.org/editions/2020

[27] Y. Morita, K. Fukuda, Y. Lin, T. Mori, W. Mizubayashi, S. Ouchi, H. Fuketa, S. Otsuka, S. Migita, and M. Masahara, "Tunnel FinFET CMOS inverter with very low short circuit current for ultralow-power Internet of Things applications," *J. Appl. Phys.*, vol. 56, pp. 1–5, Mar. 2017, doi: 10.7567/JJAP.56.04CD19.

[28] N. Thoti, Y. Li, S. R. Kola, and S. Samukawa, "Optimal intergate separation and overlapped source of multi-channel line tunnel FETs," *IEEE Open J. Nanotechnol.*, vol. 1, pp. 38–46, 2020, doi: 10.1109/ojnano.2020.2998939.

[29] N. Thoti, Y. Li, and S. R. Kola, "Scaling limitations of line TFETs at sub-8-nm technology node," in *Proc. Int. Symp. VLSI Technol., Syst. Appl. (VLSI-TSA)*, Aug. 2020, pp. 82–83, doi: 10.1109/VLSI-TSA48913.2020.9203648.

[30] N. Thoti and Y. Li, "Gate-all-around nanowire vertical tunneling FETs by ferroelectric internal voltage amplification," *Nanotechnology*, vol. 33, no. 5, Jan. 2022, Art. no. 055201, doi: 10.1088/1361-6528/ac2e26.

[31] A. Villalon, C. L. Royer, M. Cassé, D. Cooper, J.-M. Hartmann, F. Allain, C. Tabone, F. Andrieu, and S. Cristoloveanu, "Experimental investigation of the tunneling injection boosters for enhanced I_{ON} ETSOI tunnel FET," *IEEE Trans. Electron Devices*, vol. 60, no. 12, pp. 4079–4084, Nov. 2013, doi: 10.1109/TED.2013.2287610.

[32] Z. Wang, X.-W. Jiang, S.-S. Li, and L.-W. Wang, "An efficient atomistic quantum mechanical simulation on InAs band-to-band tunneling field-effect transistors," *Appl. Phys. Lett.*, vol. 104, no. 12, Mar. 2014, Art. no. 123504, doi: 10.1063/1.4869461.

[33] N. Thoti, Y. Li, and W.-L. Sung, "Significance of work function fluctuations in SiGe/Si hetero-nanosheet tunnel-FET at sub-3 nm nodes," *IEEE Trans. Electron Devices*, vol. 69, no. 1, pp. 434–438, Jan. 2022, doi: 10.1109/TED.2021.3130497.

CHANDNI AKBAR received the Doctoral degree in electrical engineering and computer engineering from the National Yang Ming Chiao Tung University, Hsinchu, Taiwan, in 2021. She is currently a Postdoctoral Researcher with the Parallel and Scientific Computing Laboratory, National Yang Ming Chiao Tung University. Her research interests include machine learning and deep learning algorithms and their applications in advanced nano-scaled semiconductor device simulation and optimization.

YIMING LI (Member, IEEE) is currently a Full Professor in electrical and computer engineering from the National Yang Ming Chiao Tung University, Hsinchu, Taiwan. He has authored or coauthored over 350 research papers appearing in international book chapters, journals, and conferences. His current research interests include computational electronics, device physics, semiconductor nanostructures, modeling and parameter extraction, biomedical and energy harvesting devices, and optimization techniques. He has been a Program Committee of IEDM, since 2011.

NARASIMHULU THOTI (Graduate Student Member, IEEE) received the B.Tech. and M.Tech. degrees from Jawaharlal Nehru Technological University, Hyderabad, India, in 2009 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Parallel and Scientific Computing Laboratory, National Yang Ming Chiao Tung University, Hsinchu, Taiwan. His research interests include the modeling and fabrication of emerging semiconductor devices. He was a recipient of the Best Poster Paper Award from IWPSD 2019, Kolkata, India.

• • •