Article

Prediction of Static Characteristic Parameters of an Insulated Gate Bipolar Transistor Using Artificial Neural Network

Qing Yao ^{1,2}, Yufeng Guo ^{1,2,*}, Bo Zhang ^{3,*}, Jing Chen ^{1,2}, Jun Zhang ^{1,2}, Maolin Zhang ^{1,2}, Xiaobo Guo ^{1,2}, Jiafei Yao ^{1,2}, Weihua Tang ^{1,2} and Jianhua Liu ^{1,2}

¹ College of Integrated Circuit Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; yqdeemail@163.com (Q.Y.); cjjcnjupt@163.com (J.C.); bravaisxx@163.com (J.Z.); zhangml5277@163.com (M.Z.); guoxb26@126.com (X.G.); jfyao@njupt.edu.cn (J.Y.); whtang@njupt.edu.cn (W.T.); jhliu_njupt@163.com (J.L.)

² National and Local Joint Engineering Laboratory for RF Integration and Micro-Packaging Technologies, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

³ State Key Laboratory of Electronic Thin Films and Integrated Devices, University of Electronic Science and Technology of China, Chengdu 610054, China

* Correspondence: yfguo@njupt.edu.cn (Y.G.); zhangbo@uestc.edu.cn (B.Z.)

Abstract: Breakdown voltage (BV), on-state voltage (V_{on}), static latch-up voltage (V_{lu}), static latch-up current density (J_{lu}), and threshold voltage (V_{th}), etc., are critical static characteristic parameters of an IGBT for researchers. V_{on} and V_{th} can characterize the conduction capability of the device, while BV , V_{lu} , and J_{lu} can help designers analyze the safe operating area (SOA) of the device and its reliability. In this paper, we propose a multi-layer artificial neural network (ANN) framework to predict these characteristic parameters. The proposed scheme can accurately fit the relationship between structural parameters and static characteristic parameters. Given the structural parameters of the device, characteristic parameters can be generated accurately and efficiently. Compared with technology computer-aided design (TCAD) simulation, the average errors of our scheme for each characteristic parameter are within 8%, especially for BV and V_{th} , while the errors are controlled within 1%, and the evaluation speed is improved more than 10^7 times. In addition, since the prediction process is mathematically a matrix operation process, there is no convergence problem, which there is in a TCAD simulation.

Keywords: artificial neural network; insulated gate bipolar transistor (IGBT); breakdown voltage; on-state voltage; static latch-up immunity; threshold voltage

Citation: Yao, Q.; Guo, Y.; Zhang, B.; Chen, J.; Zhang, J.; Zhang, M.; Guo, X.; Yao, J.; Tang, W.; Liu, J. Prediction of Static Characteristic Parameters of an Insulated Gate Bipolar Transistor Using Artificial Neural Network. *Micromachines* **2022**, *13*, 4. <https://doi.org/10.3390/mi13010004>

Academic Editor: Aiqun Liu

Received: 2 November 2021

Accepted: 20 December 2021

Published: 21 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The insulated gate bipolar transistor (IGBT) is widely used in power electronics due to its superior performance [1,2]. Analysis of its static characteristic parameters has always been a key part of the design process, given that these parameters, such as on-state voltage (V_{on}) and threshold voltage (V_{th}), can characterize the conduction capability of the device [3,4]. In addition, breakdown voltage (BV), static latch-up voltage (V_{lu}), and static latch-up current density (J_{lu}) are helpful when evaluating the safe operating area of the device and its reliability [5–7]. Conventionally, technology computer-aided design (TCAD) simulation tools are used to obtain these characteristic parameters of the device before experimental testing because of the advantage of low prediction errors [1–4]. However, this method may suffer from non-convergence when solving the semiconductor physical equations, thus decreasing the efficiency of obtaining the characteristic parameters.

Recently, machine learning techniques for predicting the electrical characteristic parameters of the semiconductor device have been booming due to their ability to learn the relationship between structural parameters and characteristics efficiently [8–13]. However, most work is limited to providing only one characteristic parameter, such as the threshold voltage of a junctionless nanowire transistor [11] or the breakdown voltage of a lateral

power device [12]. As for those that can output multiple characteristics, the characteristic parameters are extracted from the current–voltage (I - V) curve and capacitance–voltage (C - V) curve, which still relies on human-intensive operation [13]. In addition, machine learning is applied to predict the current value of IGBTs in circuits and suppress the current imbalance of parallel-connected IGBTs through artificial neural networks, or to predict the remaining lifetime of the IGBTs operating in circuits [14–16]. However, the focus of these works is on monitoring and optimizing IGBT operations in the circuit, rather than on exploring the relationship between the device’s structural parameters and characteristic parameters.

In this paper, we propose a simple multi-layer artificial neural network (ANN) framework to predict the vital static characteristic parameters of the IGBT, including BV , V_{on} , V_{th} , V_{lu} , and J_{lu} . The proposed method enables designers to predict the static characteristic parameters effectively, speeding up the device design process. After testing, for the same testing samples, the prediction speed of the ANN is improved by more than 10^7 times compared to TCAD simulation. This method avoids the non-convergence problem of TCAD simulations, and, if supported by experimental data, the method’s predictions will be closer to the experimental test values.

2. Dataset Generation and Division

Figure 1 shows the structural schematic diagram of the IGBT, and Figure 2 gives the curves of its static characteristic, including the breakdown characteristic curve, forward I - V characteristic curve, and transfer characteristic curve.

Figure 1. Schematic diagram of the IGBT structure.

Figure 2. Schematic diagram of the IGBT static characteristic curves.

As shown in the figure, the prediction targets, including BV , V_{th} , V_{on} , V_{lu} , and J_{lu} , are marked in the corresponding characteristic curves. In order to obtain the dataset of static characteristic parameters for IGBT with different structures, some vital structural parameters were generated randomly within a certain range. The detailed information on these structural parameters and their ranges is shown in Table 1. Moreover, during the data collection process, the static latch-up point was extracted from the forward I-V characteristic curve [7]. In addition, V_{on} was defined as the voltage at the anode when the anode's current density reached a certain set value (J_{set}), meaning that it could also be extracted from the forward I-V curve at a gate voltage V_G of 15 V, with J_{set} defined as 100 A/cm² [1]. We simulated the device characteristics of multiple IGBTs with the carrier lifetime of 1 μ s using the TCAD tool as the total dataset [17]. After finishing the total dataset collection from the simulation, the dataset was further divided into the training set and the testing set for the training process and the testing process of the ANN framework.

Table 1. Structural parameters and ranges of the IGBT.

Structural Parameters	Range
N-drift doping, N_d (cm ⁻³)	1×10^{13} – 1×10^{15}
N-buffer doping, N_b (cm ⁻³)	1×10^{16} – 5×10^{17}
P-well doping, N_{well} (cm ⁻³)	1×10^{17} – 5×10^{17}
P ⁺ anode doping, N_{p+} (cm ⁻³)	1×10^{17} – 1×10^{19}
N-drift thickness, T_d (μ m)	50–120
N-buffer thickness, T_{buffer} (μ m)	1–20
Channel length, L (μ m)	1–5

3. Methodology

Figure 3 shows the overall flow of the proposed approach and the designed ANN framework structure. In general, the proposed scheme can be divided into three steps: dataset processing, the training process, and the testing process. It is worth noting that, since the different input parameters have different scales, this will result in the updating of the ANN weights and biases during training being more susceptible to large-scale input parameters and the effect of input parameters with small scales on the output may be ignored. We use a normalization method to compress the structural parameters to (0, 1), and the process can eliminate the effect of unit and scale differences between input parameters in order to treat each class of input parameters equally, thereby increasing the prediction accuracy and efficiency of the ANN. Similar approaches to data pre-processing have been reported in papers related to machine learning [12,13].

Figure 3. The overall flow and structure of the proposed ANN framework.

The ANN is a nonlinear dynamic learning system, implemented using a mathematical method, which can handle complex nonlinear prediction problems by simulating biological neural networks. It consists of an input layer, hidden layers, and an output layer. The ANN structure for the proposed method is built using the Tensorflow library [18]. To obtain a robust framework for IGBT static characteristic parameter prediction, several trials and adjustments have been conducted. Finally, the hidden layers are set to four layers, and the number of neurons in each layer is 36, 20, 16, and eight, from first to last. Such an ANN structure has sufficient fitting ability for the static performance prediction task. Meanwhile, it also can avoid overfitting problems due to its extensive learning ability.

The ANN algorithm is capable of continuously optimizing the weights and biases between each layer to minimize the error between the predicted value and the actual value during the training process. We defined the error function by using the mean absolute error (MAE) loss function [11]. Moreover, we used the Adam optimizer to update the network parameters to reduce the error [19]. Mathematically, the MAE loss function is expressed as follows:

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (1)$$

where m is the number of training samples, y_i and $\hat{y}_i$ are the actual value and predicted value of the output feature of the i th training sample, respectively. Moreover, to capture the nonlinear relationship between the device structure and the static characteristic parameters, each neuron is equipped with a ReLU activation function [20].

4. Results Analysis

The following numerical results validate the proposed scheme. The prediction results of the ANN for different static characteristic parameters are first shown, then the effects of the different regression algorithms and the number of training samples on the prediction errors are investigated separately, and the reasons for the high and low average errors for different characteristics parameters are explained through physical theory. In addition, the ANN method has also shown its great ability to fit the trends of the static characteristic parameters well as they change with the device's structural parameters. Finally, we reveal the advantages in prediction speed by comparing the time of using ANN and TCAD simulation.

4.1. Prediction Results for Characteristic Parameters

Figure 4a–e are scatter plots showing the correlations between the target BV , V_{on} , V_{lu} , J_{lu} , and V_{th} and those predicted by the ANN method, while the number of training samples used for the prediction of the different characteristic parameters is described in the caption text. The number of training samples is influenced by the convergence of the simulation and the complexity of the predicted characteristic parameters, which will be discussed more closely later in this paper. Compared to the TCAD simulation results, the average errors for BV prediction and V_{th} prediction can be controlled to within 1%. For IGBT, BV falls between the avalanche breakdown and the reach-through breakdown limits, and is governed by the open-base transistor breakdown phenomenon. The open-base transistor breakdown condition for the IGBT is given by Equation (2), where the α_{pnp} is the common-base current gain, which is related to the injection efficiency γ_E , the base transport factor α_T , and the multiplication coefficient M . By derivation, α_T and M are in turn related to N_d and T_d [21]. V_{th} is primarily associated with the doping concentration of the channel region, N_{well} . The number of structural parameters affecting these two characteristic parameters is small, so the ANN can predict the BV and V_{th} of the IGBT very accurately. However, V_{on} , V_{lu} , and J_{lu} are closely related to on-resistance, and the parameters of the drift region, buffer region, and channel region all impact the on-resistance of the device. The more structural parameters affect the static characteristic parameter, the more complex the prediction of the characteristic parameter is. Nevertheless, the average errors of our proposed scheme are all within 8%, which proves that the number of layers and neurons of

the framework is well set so that it has sufficient fitting ability and can avoid overfitting problems due to the high fitting ability.

$$\alpha_{\text{PNP}} = (\gamma_E \alpha_T)_{\text{PNP}} M = 1 \quad (2)$$

Figure 4. The scatter plots of the static characteristic prediction results for the testing sets of the IGBT: (a) BV , training sample size: 7600; (b) V_{on} , training sample size: 7600; (c) V_{lu} , training sample size: 4250; (d) J_{lu} , training sample size: 4250; and (e) V_{th} , training sample size: 1400.

4.2. Comparison of Different Algorithms

To further emphasize the effectiveness of the ANN, we can compare the results of the ANN method with other predictors constructed using conventional regression algorithms, such as Gaussian process regression (GPR), support vector regression (SVR), and linear regression (LR). Table 2 shows the average errors of predictors constructed using different machine learning algorithms when predicting IGBT static characteristic parameters. The ANN achieves the most accurate predictions for each static characteristic parameter. In addition, it is worth noting that, except for the ANN, the average errors of the other three schemes for V_{lu} are large. This is because, for some structures with a very small V_{lu} , these schemes fail to capture the nonlinear condition due to the high complexity of the data distribution, so it is difficult to give accurate prediction results. However, due to the setting of the ReLU function, the multi-layer ANN we trained has a strong nonlinear fitting ability, so the prediction result is more accurate compared to the other three schemes.

Table 2. Average errors (%) of predictors constructed with different algorithms.

Characteristic Parameter	Machine Learning Algorithms			
	LR	SVR	GPR	ANN
BV	26.2	18.5	2.3	0.5
V_{on}	56.1	41.7	12.0	3.7
V_{th}	2.7	3.4	1.0	0.6
V_{lu}	993.1	540.2	321.4	7.4
J_{lu}	21.5	15.8	10.9	7.7

4.3. Effect of Sample Size on Results

We can also analyze the influence of the number of training samples on the prediction results. The testing set consisted of 400 IGBT samples with different structural parameters, and the number of samples in the training set was increased from 400 to 4000. Figure 5 shows the variation curves of the ANN's prediction average error for five static characteristic parameters as the number of samples in the training set increased.

Figure 5. Curves of average error for static characteristic parameters varying with the number of training set samples based on the ANN method.

It can be observed that the number of training set samples required to predict J_{lu} and V_{lu} are more than the other static characteristic parameters, and the average error is larger.

This is because when the latch-up effect occurs, the internal parasitic NPN transistor of the IGBT turns on and the gate loses the ability to control the current. Under such conditions, the relationship between the device's characteristics and the structural parameters is hard to capture, which assuredly increases the difficulty of the ANN's prediction task.

4.4. Prediction Results with Changing Structural Parameters

Taking the V_{on} prediction as an example, we can further analyze the fitting results of the trained ANN relating to the relationship between the structural parameters and characteristic parameters of the IGBT. Figure 6a–f show the comparisons of the prediction results of the ANN method and the TCAD simulation with different structural parameters, including N_b , N_{p+} , N_{well} , T_d , T_{buffer} , and L , respectively. Based on the initial structures, the structural parameters were sequentially changed. The initial structural parameters are shown in the caption. We have taken three different doping concentrations in the drift region into consideration.

Figure 6. (a–f) Comparison of predicted results from ANN and TCAD for the V_{on} of devices under the condition that six structural parameters N_b , N_{p+} , N_{well} , T_d , T_{buffer} , and L are changed, respectively.

As the figures show, compared with the TCAD simulation, the ANN method exhibits an accurate prediction ability in the presence of changes in the structural parameters, with the same trend as the TCAD simulation results. By comparing the points of three different colors in each figure, we can see that when other structural parameters remain unchanged, the larger N_d is, the smaller V_{on} will be. This is because the on-state resistance of the drift region decreases with the increase of N_d , and the anode voltage V_{on} required to reach a certain value of the anode current J_{set} decreases as well. In addition, larger values of the buffer region parameters N_b and T_{buffer} cause a decrease in the number of holes injected from the anode P+ region into the drift region, thereby weakening the conductance modulation effect, increasing the on-state resistance, and increasing V_{on} . An increase in the structural parameters N_{well} and L in the channel region raises the threshold voltage, causing an increase in the on-state resistance and V_{on} of the device. As the figures show, the proposed scheme can fit the effects of various structural parameters on the static characteristic parameters of the device well.

4.5. Prediction Time and Efficiency

For the prediction using the ANN method, since the testing process is mathematically a matrix calculation between the input features and the updated parameters within the ANN, the prediction time of the ANN is within 0.1 s. Yet, under the same dataset, the simulation runtime for the TCAD tool is 328,653.6 s, which does not include the time spent manually setting up the device structure. The prediction speed is therefore increased by more than 10^7 times compared with the TCAD simulation, and the convergence problem caused by unreasonable grid settings in the TCAD simulation is totally avoided, which is also beneficial for the efficiency when predicting the static characteristics parameters of the IGBTs with different structures.

Among the parameters of the initial structure, N_b is $1 \times 10^{16} \text{ cm}^{-3}$, N_{well} is $1 \times 10^{17} \text{ cm}^{-3}$, N_{p+} is $1 \times 10^{18} \text{ cm}^{-3}$, T_d is 80 μm , T_{buffer} is 10 μm , L is 2 μm , and three different cases with N_d of $1 \times 10^{14} \text{ cm}^{-3}$, $5 \times 10^{14} \text{ cm}^{-3}$, $1 \times 10^{15} \text{ cm}^{-3}$ are considered.

5. Conclusions

In this paper, we propose a multi-layer ANN predictive framework to predict multiple static characteristic parameters of an IGBT, such as BV , V_{on} , V_{lu} , J_{lu} , and V_{th} . Compared with the TCAD simulation tool, the trained ANN can achieve a speed increase of more than 10^7 times while ensuring the average errors are less than 8% when predicting the static characteristic parameters. Since the prediction process is a numerical operation of the matrix, it avoids the convergence problem that often occurs with TCAD simulations. The scheme confirms that ANN can capture the effects of changes to a device's structural parameters on a device's characteristic parameters and help designers to predict the static characteristic parameters quickly and speed up the design process. The method is extensible and when the ANN is trained with more informative datasets, the ANN will have the ability to predict the characteristics of more complexly structured devices. In addition, the method is able to predict characteristic parameters that are closer to the experimental test values based on experimental data.

Author Contributions: Conceptualization, Q.Y., Y.G., B.Z. and J.C.; methodology, Q.Y. and J.C.; software, Q.Y. and M.Z.; validation, Q.Y., J.Z. and J.Y.; resources, Q.Y., Y.G. and J.L.; data curation, Q.Y.; writing—original draft preparation, Q.Y.; writing—review and editing, Y.G., J.C. and B.Z.; visualization, Q.Y., W.T. and X.G.; supervision, Y.G., B.Z. and J.C.; funding acquisition, Y.G., J.Z. and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61904083, and 61874059, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20201206, BK20190237, and BK20211104, and in part by the Opening Project of State Key Laboratory of Electronic Thin Films and Integrated Devices under Grant KFJJ201907.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Huang, J.; Huang, H.; Chen, X.B. Simulation Study of a Low ON-State Voltage Superjunction IGBT With Self-Biased PMOS. *IEEE Trans. Electron. Devices* **2019**, *66*, 3242–3246. [\[CrossRef\]](#)
- Saito, W.; Nishizawa, S.-I. Alternated Trench-Gate IGBT for Low Loss and Suppressing Negative Gate Capacitance. *IEEE Trans. Electron. Devices* **2020**, *67*, 3285–3290. [\[CrossRef\]](#)
- Jiang, M.; Wang, Y. Numerical Study of a Thyristor Injection Insulated Gate Bipolar Transistor (TI-IGBT) Using P-N-P Collector. *IEEE J. Electron. Devices Soc.* **2019**, *7*, 979–983. [\[CrossRef\]](#)
- Luo, X.; Zhang, S.; Wei, J.; Yang, Y.; Su, W.; Fan, D.; Li, C.; Li, Z.; Zhang, B. A Low Loss and On-State Voltage Superjunction IGBT with Depletion Trench. In Proceedings of the 2020 32nd International Symposium on Power Semiconductor Devices and ICs (ISPSD), Vienna, Austria, 13–18 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 130–133.
- Duan, B.; Sun, L.; Yang, Y. Analysis of the Novel Snapback-Free LIGBT with Fast-Switching and Improved Latch-up Immunity by TCAD Simulation. *IEEE Electron. Device Lett.* **2018**, *40*, 63–66. [\[CrossRef\]](#)
- Antoniou, M.; Udrea, F.; Lophitis, N.; Corvasce, C.; De-Michieli, L. The P-Ring Trench Schottky IGBT: A Solution towards Latch-up Immunity and an Enhanced Safe-Operating Area. In Proceedings of the 2020 32nd International Symposium on Power Semiconductor Devices and ICs (ISPSD), Vienna, Austria, 13–18 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 134–137.
- Sun, L.; Duan, B.; Wang, Y.; Yang, Y. Analysis of the Fast-Switching LIGBT With Double Gates and Integrated Schottky Barrier Diode. *IEEE Trans. Electron. Devices* **2019**, *66*, 2675–2680. [\[CrossRef\]](#)
- Bankapalli, Y.S.; Wong, H.Y. TCAD Augmented Machine Learning for Semiconductor Device Failure Troubleshooting and Reverse Engineering. In Proceedings of the 2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Udine, Italy, 4–6 September 2019; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
- Raju, S.S.; Wang, B.; Mehta, K.; Xiao, M.; Zhang, Y.; Wong, H.-Y. Application of Noise to Avoid Overfitting in TCAD Augmented Machine Learning. In Proceedings of the 2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Kobe, Japan, 23 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 351–354.
- Mehta, K.; Raju, S.S.; Xiao, M.; Wang, B.; Zhang, Y.; Wong, H.Y. Improvement of TCAD Augmented Machine Learning Using Autoencoder for Semiconductor Variation Identification and Inverse Design. *IEEE Access* **2020**, *8*, 143519–143529. [\[CrossRef\]](#)
- Carrillo-Nunez, H.; Dimitrova, N.; Asenov, A.; Georgiev, V. Machine Learning Approach for Predicting the Effect of Statistical Variability in Si Junctionless Nanowire Transistors. *IEEE Electron. Device Lett.* **2019**, *40*, 1366–1369. [\[CrossRef\]](#)
- Chen, J.; Alawieh, M.B.; Lin, Y.; Zhang, M.; Zhang, J.; Guo, Y.; Pan, D.Z. Powernet: SOI Lateral Power Device Breakdown Prediction With Deep Neural Networks. *IEEE Access* **2020**, *8*, 25372–25382. [\[CrossRef\]](#)
- Mehta, K.; Wong, H.-Y. Prediction of FinFET Current-Voltage and Capacitance-Voltage Curves Using Machine Learning With Autoencoder. *IEEE Electron. Device Lett.* **2021**, *42*, 136–139. [\[CrossRef\]](#)
- Zeng, X.; Li, Z.; Gao, W.; Ren, M.; Zhang, J.; Li, Z.; Zhang, B. A Novel Virtual Sensing With Artificial Neural Network and K-Means Clustering for IGBT Current Measuring. *IEEE Trans. Ind. Electron.* **2018**, *65*, 7343–7352. [\[CrossRef\]](#)
- Zeng, X.; Li, Z.; Wan, J.; Zhang, J.; Ren, M.; Gao, W.; Li, Z.; Zhang, B. Embedded Hardware Artificial Neural Network Control for Global and Real-Time Imbalance Current Suppression of Parallel Connected IGBTs. *IEEE Trans. Ind. Electron.* **2020**, *67*, 2186–2196. [\[CrossRef\]](#)
- Ali, S.H.; Heydarzadeh, M.; Dusmez, S.; Li, X.; Kamath, A.S.; Akin, B. Lifetime Estimation of Discrete IGBT Devices Based on Gaussian Process. *IEEE Trans. Ind. Appl.* **2018**, *54*, 395–403. [\[CrossRef\]](#)
- Technology Modeling Associates. Medici: Two-dimensional Device Simulation Program, Version 2.2: User's Manual. Technology Modeling Associates, Incorporated. 1996. Available online: <https://manualzz.com/doc/23606185/device-simulation-tool> (accessed on 15 September 2021).
- Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016.
- Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, arXiv:1412.6980.
- Agarap, A.F. Deep Learning Using Rectified Linear Units (ReLU). *arXiv* **2019**, arXiv:1803.08375.
- Baliga, B.J. *Fundamentals of Power Semiconductor Devices*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010; pp. 748–751.