Supporting Information

A Study on Device Physics of Deep Ultraviolet Light Emitting Diodes

Leveraging Machine Learning

Na Lin,[a] Zhiqiang Liu,[b] Zhuoying Jiang,[a] Ying Jiang,[a] Shanshan Zhao,[a] Jinjian Yan,[b] Sijie Jiang, [a]Yikai Yun,[a] Wenjie Wei,[a] Shaoqun Li,[a] Ziang Wan,[a] Jianfeng Du,[a] Jinchai Li,[bcd] Tao Tao,[e] Kai Huang,*[bcd] Lin Li,*[a] Mengyu Chen,*[ad] Cheng Li*[ad] and Rong Zhang[bcde]

a School of Electronic Science and Engineering, Xiamen University, Xiamen 361005, P. R. China.
b Fujian Key Laboratory of Semiconductor Materials and Applications, CI Center for OSED, Department of Physics, Xiamen University, Xiamen 361005, P. R. China.
c Engineering Research Center of Micro-nano Optoelectronic Materials and Devices, Ministry of Education, Xiamen University, Fujian, Xiamen 361005, P. R. China.
d Future Display Institute of Xiamen, Fujian, Xiamen 361005, P. R. China
e Jiangsu Provincial Key Laboratory of Advanced Photonic and Electronic Materials, School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China, P. R. China.

Machine learning methods

In this work, AlGaN-based DUV LEDs grown on a c-plane sapphire substrate serve as a representative device for discussion. During data preprocessing, methods such as data missing value padding, data outlier deleting, data standardization, and data segmentation are employed. After these preprocessing, 378 effective samples recorded from published papers over past 17 years are selected for predicting LOPD, with 302 samples for training, 38 samples for validation, and the remaining 38 samples designated for testing.

Four ML algorithms are implemented with python language. The parameters are set as follows:

1. RF: we set the maximum depth to 15, the number of estimators to 100, and specified the parameters min_samples_leaf to 1 and min_samples_split to 2.
2. XGBoost: We set the maximum depth to 8, the minimum child weight to 2, the subsample ratio to 0.6, the column sample by tree to 1, and the number of estimators to 800.
3. FNN: We design a neural network that consists of three hidden layers. The first hidden layer contains 128 neurons, the second hidden layer consists of 32 neurons, and the third hidden layer comprises 8 neurons. ReLU is employed as the activation function, and the Adam optimization algorithm is selected for training. The entire model will undergo training for 2048 epochs.
4. CNN: weights are initialized with Glorot normal for 2 convolutional layers with the kernel number of 8 and 8. The kernel size is 3×3 and the dropout rate is set to 0.15. After the convolution operation, the feature map is flattened into a one-dimensional vector and subsequently processed by a fully connected layer containing 15 neurons. The epoch is set to 2500, while the ReLU activation function and Adam optimization algorithm are retained.

Table S1. Structural characteristic parameters, electrical characteristic parameters, and their corresponding abbreviations designed to predict the LOPD.

Input Features			
Feature	**Abbreviation**	**Feature**	**Abbreviation**
Thickness of the n-type AlGaN layer	N-T	Magnesium doping concentration in the first piece of the electron blocking layer	E1-dop

Aluminum content in the n-type AlGaN layer	N-Al	Thickness of the second piece of the electron blocking layer	E2-T
Silicon doping concentration in the n-type AlGaN layer	N-dop	Average aluminum content in the second piece of the electron blocking layer	E2-Al
The number of quantum well	MQW-P	Gradient of aluminum content variation from the beginning to the end of the second piece of the electron blocking layer	E2-K-Al
Total thickness of each barrier	B-T	Magnesium doping concentration in the second piece of the electron blocking layer	E2-dop
Average aluminum content in the first piece of the barrier	B1-Al	Thickness of the third piece of the electron blocking layer	E3-T
Gradient of aluminum content variation from the beginning to the end of the first piece of the barrier	B1-K-Al	Average aluminum content in the third piece of the electron blocking layer	E3-Al
Silicon doping concentration in the first piece of the barrier	B1-dop-n	Magnesium doping concentration in the third piece of the electron blocking layer	E3-dop
Thickness of the second piece of the barrier	B2-T	Number of periods in the p-type AlGaN layer	P-P
Average aluminum content in the second piece of the barrier	B2-Al	Total thickness of the p-type AlGaN layer	P-T
Gradient of aluminum content variation from the beginning to the end of the second piece of the barrier	B2-K-Al	Average aluminum content in the first piece of the p-type AlGaN layer	P1-Al
Thickness of the third piece of the barrier	B3-T	Gradient of aluminum content variation from the beginning to the end of the first piece of the p-type AlGaN layer	P1-K-Al
Average aluminum content in the third piece of the barrier	B3-Al	Magnesium doping concentration in the first piece of the p-type AlGaN layer	P1-dop
Gradient of aluminum content variation from the beginning to the end of the third piece of the barrier	B3-K-Al	Thickness of the second piece of the p-type AlGaN layer	P2-T

Total thickness of each well	W-T	Average aluminum content in the second piece of the p-type AlGaN layer	P2-Al
Average aluminum content in the first piece of the well	W1-Al	Gradient of aluminum content variation from the beginning to the end of the second piece of the p-type AlGaN layer	P2-K-Al
Gradient of aluminum content variation from the beginning to the end of the first piece of the well	W1-K-Al	Magnesium doping concentration in the second piece of the p-type AlGaN layer	P2-dop
Thickness of the second piece of the well	W2-T	Thickness of the p-type GaN layer	P-GaN-T
Average aluminum content in the second piece of the well	W2-Al	Magnesium doping concentration in the p-type GaN layer	P-GaN-dop
Number of periods in the electron blocking layer	E-P	The Shockley-Read-Hall lifetime	SRH lifetime
Total thickness of the electron blocking layer	E-T	The Auger recombination coefficient	Auger
Average aluminum content in the first piece of the electron blocking layer	E1-Al	Charge passing through a unit area per unit time	Current Density
Gradient of aluminum content variation from the beginning to the end of the first piece of the electron blocking layer	E1-K-Al		

Output Feature	
Feature	**Abbreviation**
Light output power density	LOPD

Figure S1. (a) RMSE scatter plot for the testing set at each sample size.

(b) R^2 scatter plot for the testing set at each sample size.
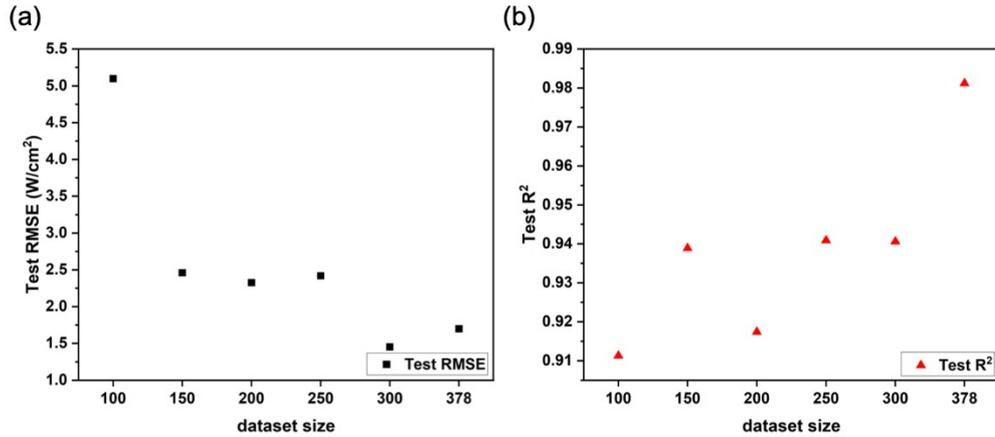
High-throughput experiments (as shown in **Figure S2a**) are conducted to investigate the variation in luminous performance of the device structure with changes in E1-Al and E1-dop (Magnesium doping concentration in the first piece of the electron blocking layer). The experimental results indicate that E1-Al is a critical feature in AlGaN-based DUV LEDs, with its underlying mechanism exhibiting significant multidimensional coupling characteristics. Specifically, an increase in the Al composition may lead to heightened lattice mismatch stress, thereby increasing the interface defect density. Moreover, excessively high barrier heights may hinder the effective injection of holes, consequently weakening the overall luminous performance. When E1-Al interacts with other structural characteristic parameters, fluctuations in SHAP values (as shown in **Figure S2b**) exhibit significant variation in both range and direction, suggesting that the influence of E1-Al is highly dependent and involves complex coupling effects. Taking E1-dop as an example, when the Al composition is higher in the samples, E1-dop typically also shows higher values. This correlation may reflect the need for a higher doping concentration at elevated Al composition to alleviate hole injection difficulties caused by increased barrier height. This interaction suggests that, in optimizing the performance of DUV LEDs, adjustments to both E1-Al and E1-dop should be considered together to improve luminous performance.

Figure S2. (a) 3D-thermal contour map of LOPD predictions with respect to E1-Al and E1-dop. (b) SHAP dependency plots for critical features in the CNN model, specifically for the features E1-Al and E1-dop.