ML-Based Fast and Accurate Performance Modeling and Prediction for High-Speed Memory Interfaces across Different Technologies

Taehoon Kim[1,2], Minjeong Kim[1,2], Hankyu Chi[2], Byungjun Kang[2], Eunji Song[2], and Woo-Seok Choi[1,*]
[1]*Dept. of ECE, ISRC, Seoul National University*, Seoul, South Korea
[2]*SK hynix*, Icheon, South Korea
{taehoon_kim, wooseokchoi}@snu.ac.kr

***Abstract*—The chip industry is undergoing a market transition from mass production to mass customization. Rapid market changes require agile responses and diversified product designs, particularly in interface circuits managing chip-to-chip communication. To facilitate these shifts, this paper proposes a machine learning-based method for rapidly and accurately predicting and analyzing the performance of high-speed transceivers, along with an evaluation methodology utilizing the proposed approach. Especially, using the process technology information as input in the dataset, this is the first work to predict the performance of a design across different technologies, which will be invaluable in architecting and optimizing designs during the early stages of development. By simulating each functional block, we gather a dataset for parameterized design and performance and incorporate device characteristics from lookup tables. The transmitter, which operates like digital circuits, is trained using parameterized signals with a DNN, while the receiver, containing analog blocks and feedback structures, employs hybrid LSTM-DNN learning with time-series input and output. Our model, trained with a 40 nm design, demonstrates high accuracy in predicting performance even with different foundries and technologies. The majority of performance parameters show an R^2 value exceeding 0.9, indicating strong predictive accuracy under varying conditions. This method provides valuable insights for early-stage design optimization and process technology scaling, offering potential for broader applications in circuit design areas.**

***Index Terms*—Memory interface, transceiver, performance prediction, machine learning, DNN, LSTM**

I. Introduction

While the semiconductor industry has traditionally focused on mass production, there has been a marked increase in demand for customized solutions tailored to unique requirements in recent years, a trend commonly referred to as mass customization [1]. In dynamic and diverse markets, rapid product planning is essential to effectively respond to specific application needs. The demand and supply of customized chips are increasing, alongside a rise in the complexity of designing interface circuits for communication between chips. However, it is challenging to find the optimal balance between data rate and energy efficiency for a target process technology at the planning stage, and achieving ambitious objectives often demands diverse circuit technologies, requiring substantial resources.

To mitigate this challenge, there have been consistent efforts to implement behavior modeling of transceivers using machine learning (ML), outperforming the accuracy and efficiency of analytical modeling. In [2-5], the nonlinearity of transceivers was successfully modeled through neural networks. However, such models can be only used for fixed link configurations, e.g., fixed channels and loads, lacking flexibility. Other prior works [6], [7] have introduced cascade-able modeling methodologies using ML, but they are all technology-specific, i.e., they cannot predict the performance of the same design in different technologies. These limitations become especially problematic in the early stage of development when circuit schemes may not be fully defined or issued process technologies may be incomplete.

Thus, in order to derive the optimal structure and design in the early stage of development, this paper proposes an ML-based method for rapidly predicting and analyzing the performance of key building blocks in high-speed wireline transceivers. This paper makes the following contributions:

- We propose a novel ML-based approach that incorporates process technology information to predict the performance of high-speed transceivers. To the best of our knowledge, this is the first work to use process technology data for predicting circuit performance across different technology nodes. This method can be utilized for process scaling predictions and to guide improvements in device fabrication, offering valuable insights for design and manufacturing optimization.
- The proposed method represents the timing performance as a parameterized signal form in order to modularize the functional blocks. This allows for flexible replacement of functional modules to optimize design and analysis. By establishing a model-independent of link configuration, it provides sufficient flexibility to adapt to various environments.
- We also present a circuit analysis environment that converts predicted signal parameters into eye-opening based on the channel, along with an application for analyzing process variation in circuits.

II. Background

A. Conventional DRAM Transceivers

Transceivers (TRx) in DRAM manage high-speed data transfer, ensuring signal integrity and reliable communication. As

This paper was result of the research project supported by SK hynix Inc.
*Corresponding author.

978-3-9826741-0-0/DATE25/© 2025 EDAA

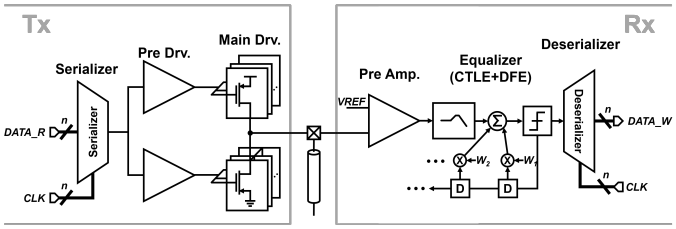

Fig. 1: Conventional high-speed TRx structure in DRAM interface

illustrated in Fig. 1, the transmitter (Tx) typically includes a serializer, pre-drivers (Pre Drv.), and a main driver (Main Drv.). During the read operation, parallel data from multiple DRAM cells are serialized into a single data stream, transferred through the pre-driver to the main driver that controls on-die termination to maintain impedance matching for signal integrity. To handle diverse channels, Tx adjusts its variable impedance value to sustain signal integrity. Process variations, voltage fluctuations, and temperature changes can affect on-resistance, requiring precise impedance control to prevent performance degradation. The receiver (Rx), shown in Fig. 1, is responsible for accurately recovering the transmitted signals, often degraded by inter-symbol interference (ISI) due to channel loss. A typical DRAM Rx data path includes a pre-amplifier (Pre Amp.) to boost the low-amplitude input signals, followed by a continuous-time linear equalizer (CTLE) to compensate for the frequency-dependent loss. Finally, a decision feedback equalizer (DFE) is employed to remove ISI by using previous symbol decisions. Incorporating both CTLE and DFE allows Rx to maintain signal integrity in high-speed environments, ensuring reliable data recovery. Both the Tx and Rx must be carefully optimized to ensure impedance matching, minimize signal distortion, and maintain low power consumption, all while adhering to the stringent area constraints of modern DRAM systems.

B. Multi-layer Perceptron

Multi-Layer Perceptrons (MLPs) are a type of feedforward neural network widely used in tasks such as classification and regression due to their ability to model complex nonlinear relationships. An MLP consists of an input layer, multiple hidden layers, and an output layer, where each layer is densely connected to the next [8]. The output y of each neuron is computed as:

$$y = \varphi(\mathbf{w}^\intercal \mathbf{x} + b), \tag{1}$$

where $\mathbf{w}$, $\mathbf{x}$, and b represent the weights, inputs, and bias, respectively. The activation function φ introduces nonlinearity into the model, which is crucial for learning complex patterns. The learning process in MLPs involves adjusting the weights and biases through backpropagation, a method where the errors of the network outputs are propagated backwards to update these parameters, minimizing the loss function. This adjustment helps improve the accuracy of prediction.

C. Long Short-Term Memory

Long Short-Term Memory (LSTM) networks are a specialized form of recurrent neural networks (RNNs) designed to address the problem of learning long-term dependencies [9]. Unlike traditional RNNs, LSTMs mitigate the vanishing gradient problem by utilizing memory cells and gating mechanisms (input, forget, and output gates) to control the flow of information. The LSTM architecture is highly effective in time-series prediction, natural language processing, and sequence generation tasks by maintaining long-term context without the degradation of gradient information during backpropagation. This makes LSTMs particularly suited for tasks where understanding the sequence and temporal relationships is critical.

D. Transfer Learning

Transfer learning (TL) is an ML technique that utilizes a pre-trained model to improve performance on a related task, especially when the new task has limited data [10]. Two common approaches in TL are: using the pre-trained model as a fixed feature extractor or fine-tuning it for the new task. In the fixed feature approach, the pre-trained model's layers are frozen, and only the final layer is adapted. Fine-tuning, on the other hand, adjusts the entire model or selected layers, making it more suitable for tasks with large differences in data distribution from the original dataset.

III. Proposed ML-based Performance Prediction

With the explosive growth in the diversity of automotive electronics, mobile, and wearable devices, the memory industry is shifting towards a market of mass customization. Despite advancements, DRAM interfaces still rely on the traditional full custom design flow, which is slow due to the complex integration of digital and analog circuits. As a result, the memory industry faces substantial challenges in managing product diversification and shortening development timelines. To address these issues, this paper proposes a method to model and predict the characteristics and performance of TRx subblocks using ML. The primary objective of our methodology is to enable rapid optimization of block configurations/designs and support decision-making in the early stage of product development and system architecture. Therefore, the methodology is based on the following core principles:

1) Standardization of Functional Block Performance: Functional block performance is standardized to allow modular signal propagation, enabling easy substitution of components with similar functions, which streamlines performance prediction and optimization across the design.
2) Technology-Independent Prediction Reliability: The method ensures reliable performance prediction even when applied to different process technologies, offering valuable insights for design decisions when transitioning between foundries or scaling down technology nodes.

This approach makes it easier to explore multiple design combinations early in development. Furthermore, by securing parameterized models for the process, it becomes possible to drive improvements in device characteristics or substantiate claims. For DRAM TRx, a comprehensive methodology that covers both the digital and analog components is necessary. The following subsections present details of data collection,

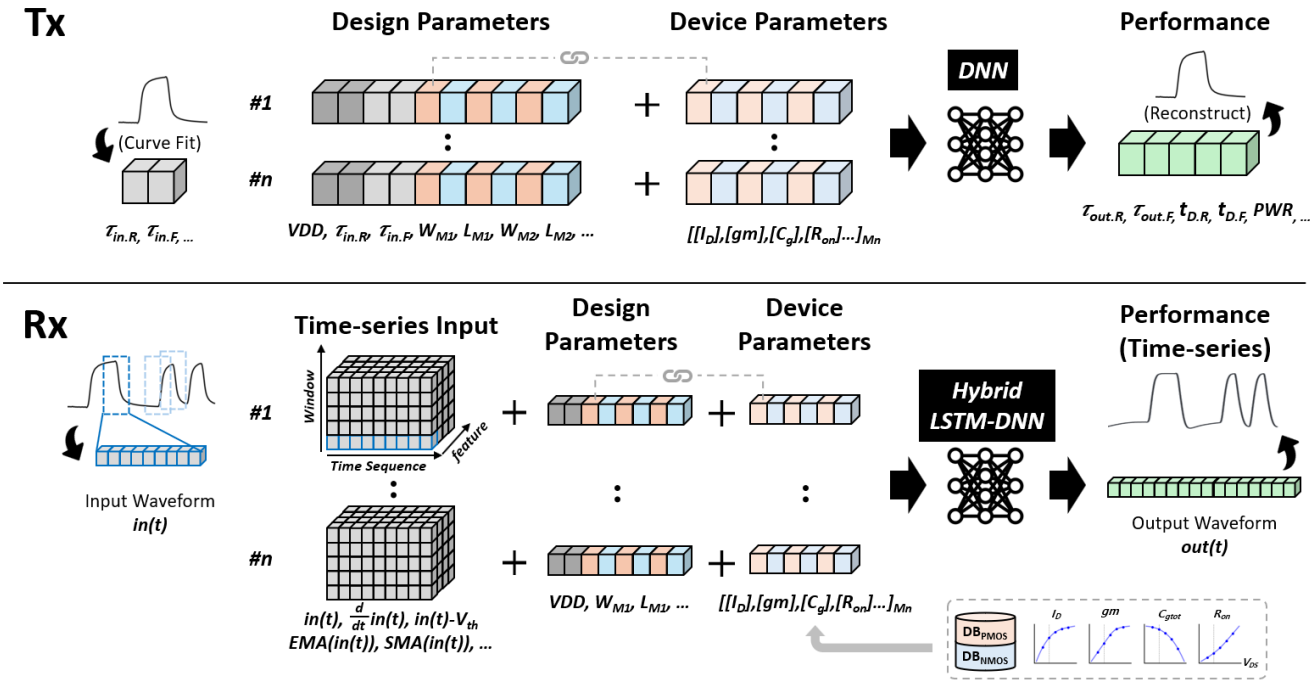

Fig. 2: Overview of the proposed ML-based performance prediction

the model training procedure, and the techniques applied for ML-based modeling.

A. ML-based Tx Performance Prediction

While this section focuses on Tx, the proposed methodology can also be effectively applied and extended to custom designs in the digital CMOS logic domain.

1) Data Preparation: The dataset is organized into three key categories: device information for various MOSFET types, design parameters, and the corresponding performance data generated from SPICE simulations based on these design parameters. Figure 2 provides an overview of the dataset configuration. The design table includes parameters such as transistor sizes, supply voltage, input stimuli time constants and delays, as well as control cases. The performance table is customizable, typically containing power consumption and output signal characteristics like time constants and delays. The device parameter table, containing the characteristics of the components used, completes the dataset. MOSFET information is pre-characterized into a lookup table (LUT) using SPICE simulations, and matched to the devices used. Characterization is performed with various process corners and device sizes, using MOSFETs with a single finger. Key device parameters, including I_D, gm, C_{gtot}, and R_{on}, are stored as functions of drain-to-source voltage V_{DS} and gate-to-source voltage V_{GS}, and importantly, all parameters are referenced to the device threshold voltage V_{th} [11]. The LUT is divided into linear and saturation regions. Characteristics are extracted at V_{DS} values of 50% and 100% of V_{th}, and at 25%, 50%, and 75% of the range between V_{th} and VDD. This approach enhances model prediction reliability across different process technologies compared to characterizing with uniformly spaced V_{DS} values.

For parameterized representation of the input and output signal waveforms, a fitting function that rises and falls in a second-order exponential manner, as defined in (2), is employed, which closely resembles the actual waveform of digital signals.

$$v(t) = \begin{cases} V_0 & t < t_{d0} \\ (V_1 - V_0) \cdot \left(1 - \exp\left(-\frac{(t-t_{d0})^2}{\tau^2}\right)\right) + V_0 & t \geq t_{d0} \end{cases} \tag{2}$$

where t_{d0} defines the time t_d, at which the signal reaches half of the voltage level when transitioning from V_0 to V_1, and is given by $t_{d0} = t_d - \sqrt{-\tau^2 \ln(0.5)}$. In this work, the input stimuli and output waveform shapes are represented using τ, t_d, V_0, and V_1, as defined in (2), and these parameters are used for training and performance prediction.

2) Model Architecture: We employ a deep neural network (DNN) model based on the MLP structure since MLPs with two hidden layers can approximate arbitrary functions [12]. The model consists of two hidden layers with widths of 1024 and 256, respectively. Each perceptron in the hidden layers utilizes the Rectified Linear Unit (ReLU) activation function, chosen for its ability to efficiently handle non-linearity and mitigate the vanishing gradient problem. The model architecture (i.e., widths and depths of MLP) is chosen through empirical evaluations.

3) Eye-opening Prediction by Post-processing MLP Outputs: Since our focus is on predicting the intrinsic performance of the Tx design, we need a post-processing method to evaluate the eye-opening under various channel conditions. While this is a traditional approach, we have refined it to be compatible with the results generated by our prediction model.

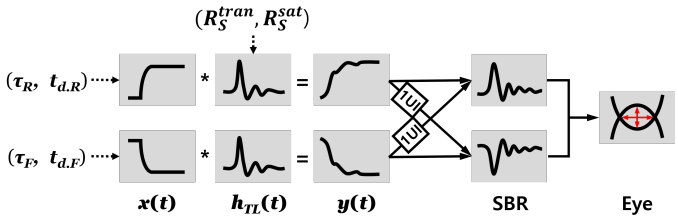

Fig. 3: Conceptual process flow of the proposed eye-opening prediction method

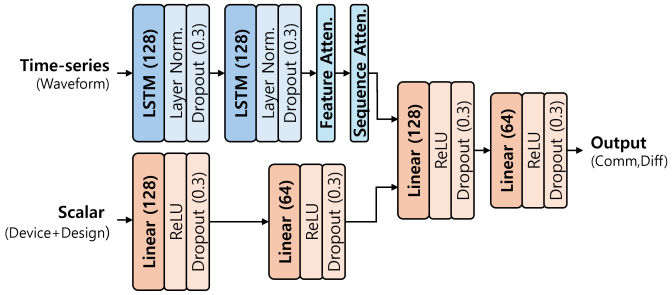

Fig. 4: Hybrid LSTM-DNN model architecture for Rx performance prediction

The proposed method calculates eye-opening under practical link environments based on the quantified output shape of the off-chip driver and source impedance parameters. This process is briefly illustrated in Fig. 3. To evaluate the eye-opening of the main driver, the model and dataset must be structured to predict the output shape parameters (τ, t_d) of the off-chip driver and the source resistance R_s^{tran} and R_s^{sat}, representing the half-VDD and full saturation conditions, respectively. Step signals $x(t)$ are generated from the signal shape parameters (τ, t_d), and are then convolved with the channel's impulse response $h_{TL}(t)$ to produce the step responses $y(t)$ at the Rx. The step response generation is handled separately for the segments when the main driver's signal is transitioning and when it approaches full saturation. In each case, the respective channel impulse response $h_{TL}(t)$ is obtained by applying the inverse Fourier transform to the transfer function $H(s)$, which includes reflection effects from the source and load sides [13].

These responses, along with their delayed waveforms, can then be manipulated to shape the single bit response (SBR). To produce a more accurate SBR, weights are applied based on the transmission line's lattice diagram, followed by the superposition of the signals. Then, using the peak distortion analysis (PDA) method [14], the worst-case eye diagram is obtained by folding the SBR waveform and its shifted versions at unit intervals.

B. ML-based Rx Performance Prediction

The proposed Tx modeling method utilizes parameterized signal waveforms for learning, which works well for logical operations but is limited when applied to analog front-ends in Rx. Furthermore, for circuits incorporating feedback like DFE, simple propagation models fail to accurately replicate the behavior. To address these challenges, we propose a hybrid LSTM-DNN approach for predicting Rx performance. While the dataset is still constructed using design and device parameters, the key distinction is that the model is trained on time-series input waveforms rather than parameterized stimuli, allowing for direct prediction of the output waveform.

1) Data Preparation: The dataset for Rx consists of both time-series data and scalar data. The scalar dataset is constructed similarly to the Tx methodology, containing design and device parameters, but without the input stimuli. Time-series input and output waveforms are obtained from SPICE simulations corresponding to the design parameters and organized into batches based on user-defined time windows. Each batch is paired with the corresponding design and device parameters for training. To enhance model generalization, the input stimuli generated during SPICE simulations have various swing range, slew rate, and slight frequency fluctuations to create a more diverse dataset. The common and differential mode voltages for the input and output of each function block are defined as a standardized format. To account for the nonlinear characteristics and hysteresis of the analog front-end, the input waveforms are preprocessed to generate additional input features. These features include the waveform's derivatives, simple moving average (SMA), and exponential moving average (EMA). To account for feedback from previous symbols and the effects of ISI, the maximum frame size for the moving averages is set beyond these intervals to ensure accurate modeling. The aim of incorporating past information in the learning process is to effectively capture time-invariant characteristics, ensuring more accurate predictions.

2) Model Architecture: We propose a hybrid learning approach that predicts the output waveform from the input waveform and its derived features, while incorporating design and device parameters into the model. To capture the temporal dependencies and memory effects of the Rx circuits, LSTM layers are employed, while DNNs are used to model the system nonlinearity. The model consists of two LSTM layers, with layer normalization to enhance training stability and dropout to prevent overfitting and improve generalization. Attention layers are applied in both the feature and sequence dimensions, enabling the model to focus on the most relevant information and boosting performance on time-series data. Both scalar data processing and integration of the LSTM outputs are managed through DNNs, structured as MLPs with ReLU activation functions and dropout layers to ensure stable learning and regularization. The proposed architecture is depicted in Fig. 4.

C. Transfer Learning

To improve accuracy, the pretrained model is fine-tuned using transfer learning. The training set for fine-tuning is derived from arbitrarily chosen design parameters, which can be easily obtained through technology porting by leveraging existing schematics. Since this approach spans different process technologies, the dataset exhibits low similarity in process information, and applying fine-tuning is effective to improve the model performance, as will be shown in Section IV.

TABLE I: MODEL TRAINING CONFIGURATION

	Tx (DNN)	Rx (LSTM+DNN)
Training set (Design set)	20000 (Pre-trained) 100 (Fine-tune)	5000 (Pre-trained) 100 (Fine-tune)
Data usage ratio	100%	10%
Test set	20000	500
Stimuli for dataset	Step	PRBS
Epochs	400	150
Loss Function	MAE (L1)	Smooth L1
Training time	$<$ 5 min	$<$ 10 hours

IV. EXPERIMENTAL RESULTS

In this section, we demonstrate the accuracy and effectiveness of the trained models for the proposed methodology. The method has been applied to the DRAM TRx circuits as described in II-A. The implemented interface operates at a data rate of 10 Gbps with a clock frequency of 2.5 GHz. The dataset is generated using Synopsys HSPICE simulations across three process technologies provided by two foundries. The following two cases are considered:

- Case 1: From 40 nm to 28 nm technology on Foundry A
- Case 2: From 40 nm technology on Foundry A to 28 nm technology on Foundry B.

The design parameters in the training set are randomly chosen and includes factors such as supply voltage, skew corner, stimuli shape, gate sizes, load transistor sizes, and numerically represented control cases. The settings applied during training are summarized in Table I. The training time, excluding data gathering, is listed in the table. As it can vary depending on the environment, the maximum time is reported.

A. Tx Performance Prediction

The proposed method is evaluated for three key Tx components. Table II presents the mean absolute error (MAE), mean relative error (MRE), and the coefficient of determination (R^2) for performance, as predicted by the model trained on Foundry A's 40 nm technology for two different 28 nm technologies. For Case 1 and Case 2, MAE for signaling performance metrics, such as τ and t_d, becomes at most 0.55 ps and 1.25 ps, respectively. Power consumption is predicted for internal blocks only, with a maximum MRE of 2.53 %, demonstrating the model's strong predictive accuracy. Overall, the model demonstrates strong performance with R^2 values exceeding 0.9 in the majority of cases. Among the predicted parameters, the delay parameter distribution is displayed in Fig. 5 as delay exhibits significant variation depending on the technology. The predicted values closely align with the actual values, and it is noteworthy that although the training and test sets display different distributions, the model maintains strong prediction reliability across different technologies.

In Section III-A3, a post-processing methodology to determine the eye-opening based on the predicted signaling parameters of the main driver is introduced. We demonstrate the effectiveness of this approach. For Case 1, the post-processed results derived from the predicted parameters are compared with SPICE simulation results for the corresponding circuit. As shown in Fig. 6, a scatter plot compares the timing (width) and voltage (height) margins of the eye-opening, derived from the predicted parameters, with those from the reference eye-opening. The R^2 values for the timing and voltage margins are 0.97 and 0.99, respectively, indicating a strong linear correlation between the predicted and reference results. The MAE of 1.63 ps for timing and 7 mV for voltage margins demonstrate sufficient accuracy for evaluating design robustness.

TABLE II: PREDICTION RESULTS OF SUB-BLOCKS IN TX

Block	Target Param.	Case 1 A 40nm - A 28nm			Case 2 A 40nm - B 28nm		
		MAE	MRE[%]	R^2	MAE	MRE[%]	R^2
Serializer	τ_R [ps]	0.23	3.32	0.95	0.38	4.25	0.93
	$t_{d.R}$ [ps]	0.49	0.86	0.99	1.25	1.43	0.97
	Power [uW]	0.69	0.8	0.99	0.56	0.69	0.99
PreDrv. Unit	τ_R [ps]	0.14	5.4	0.87	0.26	5.49	0.85
	$t_{d.R}$ [ps]	0.32	2.19	0.99	0.62	2.34	0.98
	Power [uW]	0.21	2.53	0.99	0.19	2.33	0.99
Main Drv.	τ_R [ps]	0.55	11.78	0.8	0.55	9.97	0.86
	$t_{d.R}$ [ps]	0.39	0.36	0.91	0.51	0.47	0.90
	Rs^{sat} [Ω]	0.52	0.89	1.00	0.95	1.83	0.99

Additionally, the proposed model can be used for process variation analysis in circuits as follows. When attaching device parameters from the LUT, components with random variations generated through Monte Carlo (MC) simulations are used as the model inputs. We perform the process variation analysis using a model trained on Foundry A's 40 nm technology. The device LUT is constructed using 1,000 MC-simulated MOSFET samples with local variations at the typical corner on Foundry A's 28 nm technology. For comparison, the reference data are obtained by simulating 1,000 samples using HSPICE Monte Carlo on the same design set. Figure 7 illustrates the histograms of the rise edge delay for the serializer and pre-driver, along with predictions utilizing the LUT with variations. The mean and standard deviation differ by less than 1 ps, demonstrating the proposed method's accuracy in analyzing process variations.

B. Rx Performance Prediction

The experiments are conducted on key components of a conventional Rx analog front-end, including the pre-amplifier, CTLE, and DFE, to demonstrate the effectiveness of the hybrid learning method in predicting their performance. Since the time-series prediction directly generates the output waveform,

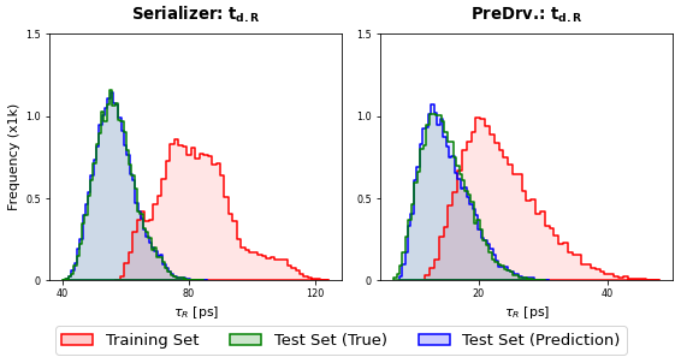

Fig. 5: Data distribution of training and test sets

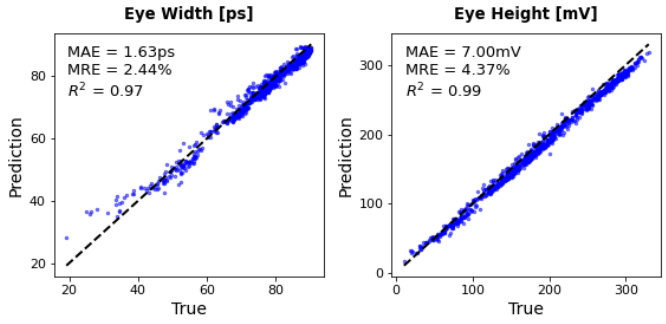

Fig. 6: Comparison of eye margins between post-processed predicted parameters and reference simulations

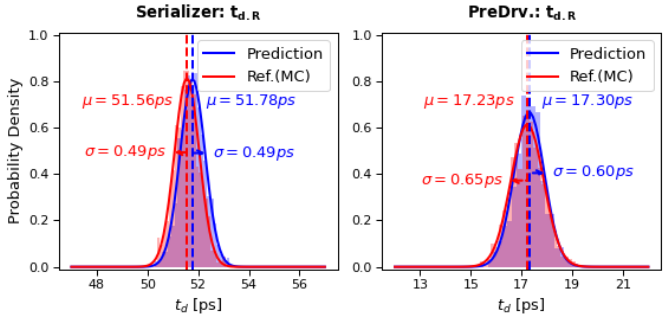

Fig. 7: Application of the proposed methodology for variation analysis

the accuracy of the ML model is evaluated using the commonly employed eye-opening margin in the field. Table III presents the errors and R^2 values for the timing and voltage margins at the output nodes of each functional block. For the pre-amplifier and CTLE, the MRE is below 5 %, and the R^2 exceeded 0.97, indicating a high level of accuracy in matching the actual sample's margins. While the DFE shows slightly lower accuracy, the model is still considered suitable for use in the early stages of architecture development. Figure 8 shows the output waveform of the CTLE for a selected design from the test set. All design parameters and input stimuli are kept identical, except for the device parameters. The ML-based prediction closely aligns with the SPICE simulation, and more importantly, the model accurately reflects the technology information.

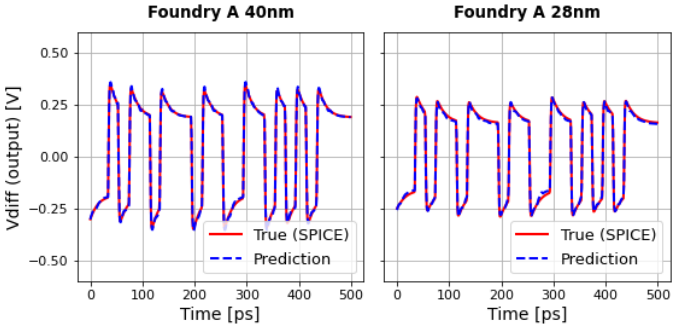

Fig. 8: CTLE output waveform accuracy and model reliability across different technologies

TABLE III: PREDICTION RESULTS OF SUB-BLOCKS IN RX

Block	Performance	Case 1: A 40nm - A 28nm		
		MAE	MRE[%]	R^2
PreAmp.	Eye Width [ps]	1.09	1.53	0.98
	Eye Height [mV]	16.40	3.89	0.98
	Power [uW]	33.41	3.72	0.97
CTLE	Eye Width [ps]	0.51	0.69	1.00
	Eye Height [mV]	14.38	4.10	0.99
	Power [uW]	44.60	3.98	0.98
DFE	Eye Width [ps]	2.26	3.45	0.86
	Eye Height [mV]	28.38	16.12	0.88
	Power [uW]	24.08	3.61	0.98

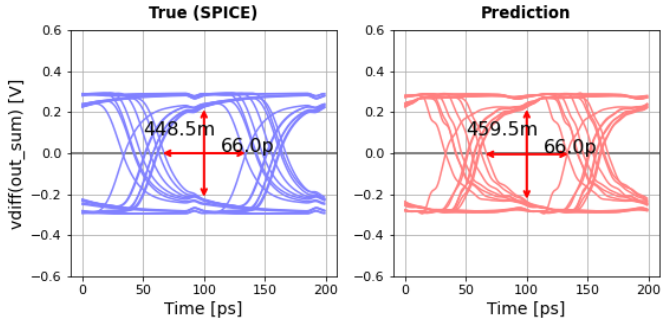

Fig. 9: Comparison of DFE summation block output waveform between ML-based prediction and SPICE simulation

In Fig. 9, the output waveform of the summing node of the DFE is displayed. The predicted results accurately track the eye height variations based on the DFE tap coefficients, closely matching the reference result. This demonstrates that the proposed methodology provides reliable and technology-independent performance prediction.

V. CONCLUSION

This paper proposes an ML-based method for quickly and accurately predicting and analyzing the performance of high-speed transceivers. For each functional block, we collect design parameters, device information, and corresponding performance as a dataset, and train performance predicting models using both DNN and a hybrid LSTM-DNN architectures. As a technology-independent model, our approach enables high-accuracy performance prediction across different technology nodes, making it a valuable tool for early-stage design optimization and process shrinkage analysis. We validate the effectiveness of the proposed method by evaluating the accuracy of performance prediction across different technologies. When applying the model trained on one technology node to a different foundry's technology, the accuracy of the predicted performance remains highly accurate.

In conclusion, the proposed ML-based approach offers a robust and adaptable framework for performance prediction and analysis in high-speed wireline transceivers. Its flexibility, accuracy, and technology-independent characteristics make it a promising tool for future circuit design and optimization across various process technologies.

ACKNOWLEDGEMENT

The EDA tool was supported by the IC Design Education Center (IDEC), Korea.

REFERENCES

[1] S. Davis. (2023) Reshaping the semiconductor landscape – generative ai, customized chips, and memory outlook in 2024. [Online]. Available: https://www.semiconductor-digest.com/reshaping-the-semiconductor-landscape-generative-ai-customized-chips-and-memory-outlook-in-2024/

[2] T. Lu, J. Sun, K. Wu, and Z. Yang, "High-speed channel modeling with machine learning methods for signal integrity analysis," *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 6, pp. 1957–1964, 2018.

[3] X. Chu, J. Li, X. Li, J. Wang, and Y. Li, "Modeling for nonlinear high-speed links based on deep learning method," in *2019 12th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo)*, 2019, pp. 19–21.

[4] X. Wang, T. Nguyen, and J. E. Schutt-Ainé, "Laguerre–volterra feed-forward neural network for modeling pam-4 high-speed links," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 10, no. 12, pp. 2061–2071, 2020.

[5] Y. Cao, I. Erdin, and Q.-J. Zhang, "Transient behavioral modeling of nonlinear i/o drivers combining neural networks and equivalent circuits," *IEEE Microwave and Wireless Components Letters*, vol. 20, no. 12, pp. 645–647, 2010.

[6] Y. Zhao, T. Nguyen, H. Ma, E.-P. Li, A. Cangellaris, and J. E. Schutt-Ainé, "Modeling cascade-able transceiver blocks with neural network for high speed link simulation," in *2022 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)*, 2022, pp. 1–3.

[7] Y. Zhao, T. Nguyen, H. Ma, E.-P. Li, A. C. Cangellaris, and J. E. Schutt-Ainé, "Modular neural network-based models of high-speed link transceivers," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 13, no. 10, pp. 1603–1612, 2023.

[8] S. Haykin, *Neural Networks and Learning Machines*, ser. Pearson International Edition. Pearson, 2009. [Online]. Available: https://books.google.co.kr/books?id=KCwWOAAACAAJ

[9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[10] R. Liu, Y. Shi, C. Ji, and M. Jia, "A survey of sentiment analysis based on transfer learning," *IEEE Access*, vol. 7, pp. 85 401–85 412, 2019.

[11] M. Kim, T. Kim, and W.-S. Choi, "Device parameter extraction method for training performance predicting models," in *2024 21st International SoC Design Conference (ISOCC)*, 2024, pp. 424–425.

[12] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/089360809190009T

[13] T. Esmailian, F. R. Kschischang, and P. G. Gulak, "In-building power lines as high-speed communication channels: channel characterization and a test channel ensemble," *International Journal of Communication Systems*, vol. 16, 2003. [Online]. Available: https://api.semanticscholar.org/CorpusID:10726141

[14] B. Casper, M. Haycock, and R. Mooney, "An accurate and efficient analysis method for multi-gb/s chip-to-chip signaling schemes," in *2002 Symposium on VLSI Circuits. Digest of Technical Papers (Cat. No.02CH37302)*, 2002, pp. 54–57.