IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 December 2023, accepted 2 January 2024, date of publication 9 January 2024,
date of current version 18 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3351843

RESEARCH ARTICLE

Length-Dependent Deep Neural Network Based Modeling for High-Speed Channels

HUNG KHAC LE[1], (Student Member, IEEE), AND SOYOUNG KIM[2], (Senior Member, IEEE)
[1]Department of Electrical and Computer Engineering, College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, South Korea
[2]College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, South Korea

Corresponding author: SoYoung Kim (ksyoung@skku.edu)

This work was supported in part by Samsung Electronics Co., Ltd under Contract ID MEM210728_0001, and in part by the National Research Foundation of Korea funded by the Ministry of Science and Information and Communication Technology (ICT) under Grant 2020R1A5A1019649 and Grant 2020M3H2A1076786.

ABSTRACT This article presents a length-dependent deep neural network (LD-DNN) based channel modeling methodology to predict the frequency response of high-speed channels. The proposed method significantly enhances the model accuracy and design efficiency while considering the channel length dependence that was neglected in previous modeling approaches. We define the concept of the electrical length to model the length and frequency dependence, then further leverage the activation function to capture the multiple reflection effects to improve accuracy. Additionally, we model the insertion loss resonance induced by crosstalk that can seriously deteriorate signal integrity. As a result, by adopting the proposed model which can predict the S-parameters as a function of length, the need for performing additionally 3D electromagnetic simulations when adjusting the channel length can be eliminated. Various high-speed channel cases are tested to validate the accuracy of the proposed method. The modeling accuracy is less than 4% for different high-speed channel structures with run times of less than 1.4 second per design.

INDEX TERMS Activation function, deep neural network (DNN), electrical length, high-speed channel, multiple reflection, resonance, S-parameter, transmission line.

I. INTRODUCTION

Machine learning is a powerful technique for modeling electromagnetic (EM) response [1], [2], [3], [4]. Recently, machine learning has been more involved in the plurality of electronic design processes due to the breakthrough of deep learning [5]. In particular, reinforcement learning algorithms are well-suited to decision-making tasks such as circuit design [6], [7], [8]. In addition, optimization procedures widely employ active-learning algorithms such as Bayesian optimization [9], [10], [11]. Alternatively, supervised and unsupervised learning are extensively exploited to enhance design efficiency using machine learning-based surrogate models [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. To this end, contemporary deep learning techniques with breakthroughs in various applications, such as deep neural networks (DNNs), have emerged as candidates for this surrogate model. For 3-D components like high-speed channels, the design process requires optimizing a structure in the frequency domain before performing time-domain verification. Thus, an efficient design technique in the frequency-domain or time-domain can improve the interconnect design procedure. For example, various authors [12], [13], [14] have attempted to enhance the time domain procedure by predicting the eye diagram. While the proposed approaches in [12] and [13] provide a fast solution to estimate eye height and width using the DNN model trained by frequency response data, the others [14] leveraged feature selection techniques depicted in [32] and [33] to improve the DNN model performance in [13]. Additionally, the authors in [15] utilized the DNN model to forecast crosstalk in high-speed transmission lines. However, these approaches require training data from both frequency-domain and time-domain simulations, resulting in inefficiencies due to the complicated structure.

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

High-frequency design characterization is one of the most time-consuming design procedures because of 3-D electromagnetic (EM) simulations. As a result, the typical strategy is to employ a surrogate model to produce reliable data in the frequency domain instead of repetitively performing EM simulations. Although the DNN application in a high-speed channel design has receives relatively little interest [26], [27], [28], DNN-based microwave component modeling has attracted significant attention [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. Nonetheless, it is noteworthy that these applications share the standard approach to capture the frequency response for a given structure geometry and frequency range [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. An alternative approach with a neuro-based transfer function that includes two approximation processes to model the frequency response has been reported to model frequency response [16], [17], [18]. First, this approach transforms the frequency response (such as S-parameters provided by the EM simulator) into pole-zero-gain transfer functions. Then, a neural network discovers the relationship between geometrical variables and these transfer functions. As a result, the transfer functions predicted by DNN incorporate frequency information to estimate the frequency response. Although this approach eliminates the frequency in the process for efficient learning, multiple approximation processes can introduce significant errors. On the other hand, a knowledge-based method described in [29], [30], and [31] utilizes the knowledge in existing empirical/equivalent circuit models with neural networks to efficiently develop a more accurate model. However, the quality of empirical models somehow confines the ultimate prediction accuracy. Additionally, serial modeling processes are more likely to produce more errors than direct modeling. Furthermore, empirical models are not always available, and it also takes significant effort to optimize a new equivalent circuit before employing the knowledge-based technique.

Dong et al. [19] proposed using DNN for multi-objective optimization for a multi-parameter antenna structure using an ordinary L1 regularization technique that punishes the cost function to prevent overfitting. Alternatively, Jin et al. [20] modified the DNN topology with two stages. The authors considered that the impact on the EM response caused by the geometrical parameters is more complex than the impact produced by frequency. Therefore, the authors allocated more hidden layers for the first stage to learn the relationship between geometrical variables and EM response. Then, the output of the first stage concatenates with the frequency to produce the final EM response with fewer hidden layers in the second stage. In contrast, Jin et al. [21] focused on the activation function leveraged in the DNN model to capture the non-linear components in the frequency response. In this work, the authors proposed a modification of the ReLU function, referred to as the smooth ReLU (SReLU) function, which captures the non-linear components of frequency response and avoids the vanishing gradient problems in the neural network. Similarly, Sun et al. [22] introduced a

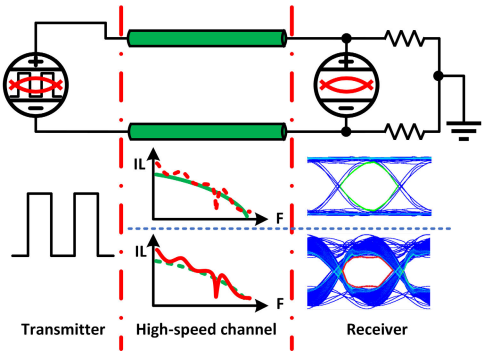

FIGURE 1. Influence of reflections and resonance on the receiver eye diagram.

DNN-based design approach for antenna structure design. Another application described in [23] and [24] involves modifying the DNN model with specific additional layers to ensure the causality and passivity of the predicted frequency response. Alternatively, Torun et al. [25] attempted to use a 1-D convolutional neural network to preserve the spatial relationship in the frequency axis of a solenoidal inductor.

While the techniques mentioned above focus on antenna design, other approaches [26], [27], [28] utilized DNNs for high-speed channel design. The method proposed in [26] and [28] involves predicting the RLGC components of the high speed channel extracted from a 2D extractor based on the cross section and then using these components to estimate S-parameters. However, this approach introduces two significant drawbacks in high-speed channel modeling. The data obtained from the 2D extractor are only qualified for initial analysis and cannot capture various effects caused by channel length, which are only observable through 3D full-wave simulations. Secondly, this technique is incapable of simulating complex structures that require 3D simulations such as serpentine lines, making it inconvenient for intricate designs. Alternatively, Le and Kim [27] leveraged DNN as a surrogate model for channel frequency response. Then, an optimization algorithm employs this black-box model to obtain the best design structure for saving power. While frequency response is not the ultimate metric for evaluating a high-speed channel, it is a mandatory component for selecting the most appropriate design. When employing a DNN-based surrogate model, improper frequency response predictions can lead to undesirable scenarios in design optimization, as shown in Fig. 1. The failure to identify the resonance frequencies damages the eye diagram at the receiver by a significant drop in the eye height and width. Besides, the undetected reflections either constructively or destructively influence the eye diagram at the receiver due to amplitude distortion in the signal. Furthermore, inaccurate phase response can also result in improper inter-symbol interference, which leads to a wrong eye diagram at the receiver and inappropriate equalization solutions. As a result, the DNN-based surrogate models should provide reliable numerical forecasts and identify reflections or resonance

effects in the high-speed channel to thoroughly replicate the EM simulator.

In this paper, we propose a length-dependent deep neural network (LD-DNN)-based modeling technique to address the limitations of high-speed channel frequency response prediction. While existing methods originated from machine learning perspectives that assume the intricacy of an DNN structure can mimic any complex function, our proposed approach utilizes valuable information extracted from the knowledge domain instead of increasing model complexity for efficient and generalized learning. Accordingly, the proposed technique presents a comprehensive accuracy enhancement over existing methods. Our approach introduces an informative input to improve LD-DNN model performance. Furthermore, the selected activation function and resonant feature modeling based on the signal transmission mechanism are employed to reveal all phenomena in the frequency response comprising multiple reflections and resonance. Moreover, the proposed method can quickly predict channel frequency response when adjusting the channel length, which requires hours of simulation using an EM simulator. It also addresses the length dependence that was not considered in the existing surrogate model approaches for high-speed channels [26], [27], and [28]. The rest of the paper is organized as follows: Section II describes the background theory inspiring the proposed methodology, including information about the electrical length, multiple reflection effects, and resonance in the insertion loss. Section III proves the effectiveness of using the electrical length as an input feature. The activation function selection and resonant feature modeling are also discussed in this section. Section IV describes the numerical results for various high-speed channel structures with the phase and magnitude of the S-parameter prediction. Finally, the conclusions of this paper are given in Section V.

II. BACKGROUND THEORY

The surrogate model for the high speed channel can remove the process of running 3D EM simulations during the design stage, which can significantly reduce the number of design iterations needed to find the optimal design. Consequently, the surrogate model should accurately predict the frequency-dependent channel response. Hence, we will review the transmission line theory that we used to develop comprehensive modeling techniques. We clarify the multiple reflection impact caused by the impedance mismatch and the effects of resonance or suck-out caused by crosstalk in a high-speed channel.

A. TRANSMISSION LINE THEORY AND MULTIPLE REFLECTIONS

Fig. 2 shows the signal propagation on a transmission line. According to theory, at any point (z) on the transmission line, there is an incident voltage (V_z^+) and a reflected voltage (V_z^-). As shown in (1), the total voltage at location z is the

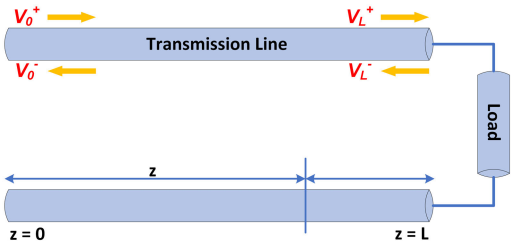

FIGURE 2. Signal propagation on a transmission line.

superposition of the incident and reflected voltages.

$$\begin{aligned} V_s(z) &= V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z} \\ \gamma &= \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \\ \beta &= \frac{2\pi}{\lambda} = \frac{2\pi f \sqrt{\varepsilon_r}}{c} = kf \end{aligned} \tag{1}$$

where:

z: The location with respect to source
γ : Propagation constant
α : Attenuation constant
β : Phase constant
λ : Wavelength
c : Speed of light
ε_r : Relative permittivity
f : Frequency

The signal in the transmission line depends on the phase constant β and its location z as shown in (1). As a result, the signal at the load or receiver end is represented by (2).

$$V_s(L_p) = V_0^+ (e^{-(\alpha+j\beta)L_p} + \tau_L e^{-(\alpha+j\beta)L_p}) \tag{2}$$

where:

τ_L : Reflection coefficient at load
L_p : physical channel length

The signal strength is generally inversely proportional to the frequency in an ideal case without impedance mismatch. Therefore, the curvature is flat and there is no reflection. However, achieving an optimal situation in a wide range of experimental designs (DoE) is difficult. There are always mismatched impedance structures in this region. Multiple reflections occur when the physical channel length is significantly greater than the wave length. As a result, the mechanism behind this occurrence is depicted in Fig. 3(a) as a function of the physical channel length.

In Fig. 3(a), the orange signal represents the incident signal and the blue one represents the reflected signal. The transmission line with Z_0 impedance connects the driver on the left side and the receiver on the right side. When the physical channel length equals a quarter of the wave length ($\frac{\lambda}{4}$), the return loss reaches a local maximum while the insertion loss achieves a local minimum. First, the incident signal from outside encounters the mismatched impedance between the transmission line and driver termination. As a

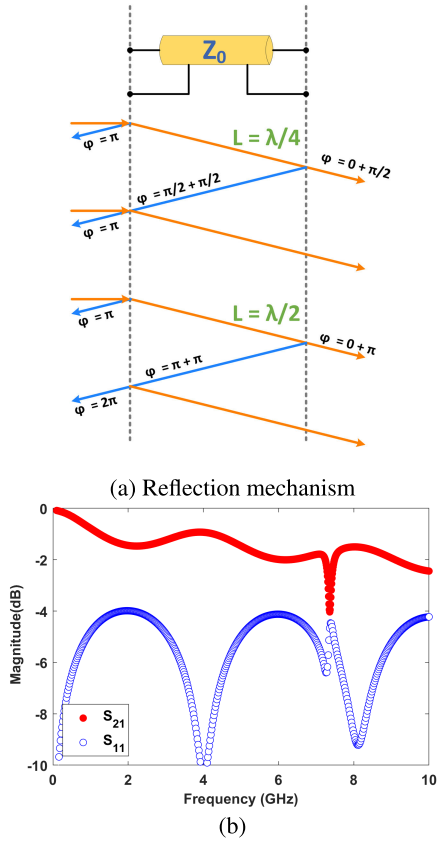

FIGURE 3. Multiple reflections. (a) Mechanism and (b) impacts on insertion loss and return loss.

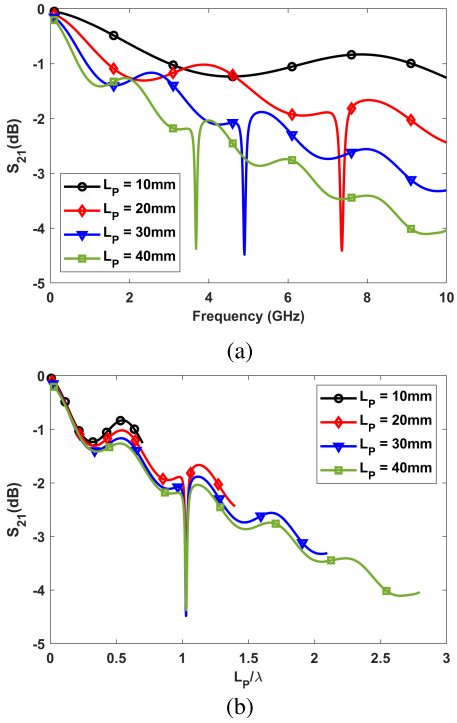

FIGURE 4. Resonance effects versus (a) frequency (b) electrical length (L_e).

result, a part of the incident signal keeps traversing along the transmission line to reach the receiver. Meanwhile, the remaining portion is directly reflected. Hence, this reflected signal is out of phase with the incident signal (π). Additionally, the transmitted signal suffers a phase delay of $\frac{\pi}{2}$ before encountering an impedance mismatch at the receiver. A portion of this signal is still being transmitted. Meanwhile, the remaining signal reflects toward the driver. This reflected signal also suffers a phase delay of $\frac{\pi}{2}$ caused by the channel length, resulting in a total phase delay of π. As a result, the final reflection signal is the sum of the two previously mentioned in-phase reflected signals. Therefore, this constructs a local maximum in the reflected signal corresponding to a local minimum in the transmitted signal.

The 3D EM simulators also account for this effect when the simulation results indicate that the signal moves back and forth between the driver and receiver as illustrated in Fig. 3(b). The local minimum in the insertion loss corresponds to the local maximum in the return loss and vice versa. This result supports the reflection mechanism explained above.

B. RESONANCE IN INSERTION LOSS INDUCED BY CROSSTALK

It is well known that the resonance effect in insertion loss is a severe problem in interconnect design. This phenomenon is also known as a "suck outs," "dropouts," "dips," or "notches" [35], [36]. Several causes of the resonance effect in high-speed channels are now well understood. The resonance may be the consequence of the difference in the signal skew or a stub in the channel. The resonance can be caused by the far-end crosstalk (FEXT) and also from an unusual design structure. However, the suck-out caused by crosstalk can be discussed as a general model.

In the frequency domain, resonance can occur at various frequencies for a given channel length throughout a broad frequency range, as shown in Fig. 4(a). Let's define the ratio between the physical length (L_p) and wavelength (λ) to characterize the phase shift introduced in the signal. For a constant L_p, each frequency point corresponds to a unique value of this ratio. Therefore, these resonances happen at specific ratios when we convert the frequency axis to $\frac{L_p}{\lambda}$ as shown in Fig. 4(b). This resonance is a consequence of the transmission line theory summarized in (2). The signal at the receiver end in the theory is shown in the following equation:

$$|V_s(L_p)| = |V_0^+||1 + \tau_L||e^{-\alpha L_p}||e^{-j\beta L_p}| \tag{3}$$

For simplification, we assume that the reflection coefficient phase is zero, where the load only contained the resistor but no inductance or capacitance. This is reasonable because the load is typically stable, and the reflection coefficient phase is independent of frequency or channel length. According to (3), the receiver signal will be in-phase with the incident signal when its phase component is an integer of 2π. In this scenario, if the stimulus signal in the

primary channel and the aggressor are both out of phase, FEXT will deliver an out-of-phase signal with the transmitted signal. As a result, this induced signal will be subtracted directly from the transmitted signal causing the insertion loss resonance. The simulation results, as shown in Fig. 4, strongly supported this. Equation (4) below describes the condition when the resonance effect happens.

$$\beta L_p = 2\pi N \implies \frac{L_p}{\lambda} = N \tag{4}$$

where:

N : Integer number

We define, the electrical length, L_e, as shown in (5). This L_e term is directly proportional to the ratio of L_p to λ. While the physical length in the input data, L_p, is usually apparent, this quantity is hidden but critical in the channel frequency response.

$$L_e = \frac{L_p}{\lambda} \tag{5}$$

The proposed methodology using the concept of electrical length, L_e, is described in the following section.

III. PROPOSED METHODOLOGY

The proposed LD-DNN for channel frequency response modeling with channel length variation is described in this section. First, we propose using the electrical length as an input in our LD-DNN model. Second, we propose using the snake activation function to better model the multiple reflection effects in the transmission line. Third, we develop a resonant feature to predict an essential signal integrity issue known as resonance in insertion loss. These behaviors are all related to the channel length, whose variation has been ignored in previously published work. As a result, the proposed method improves the modeling performance by leveraging domain knowledge rather than DNN structure complexity. Finally, we present the complete design methodology.

A. PROPOSED DNN STRUCTURE WITH ELECTRICAL LENGTH AS AN INPUT

Common approaches to improve the model performance of DNN include increasing the neural network complexity or improving information quality. For the former case, the model performance usually increases with an increasing number of neurons, and hidden layers, as well as an appropriate activation function. For the latter, the model performance can be drastically boosted by increasing the training data volume or providing accurate inputs to eliminate the information noise. According to the universal approximation theorem, a neural network structure can approximate any complex function. This is only true if we can provides sufficient data with appropriate inputs. It follows from transmission line theory that both L_p and L_e supply essential information to the model in the learning process. However, in previous approaches, L_e was bypassed in characterizing the channel frequency response. Suppose the conventional neural network in Fig. 5 represented by $h(x)$ predicts the frequency response from the input geometry and physical channel length ($\vec{x}$), and frequency input (f).

$$y = h(\vec{x}, f) \tag{6}$$

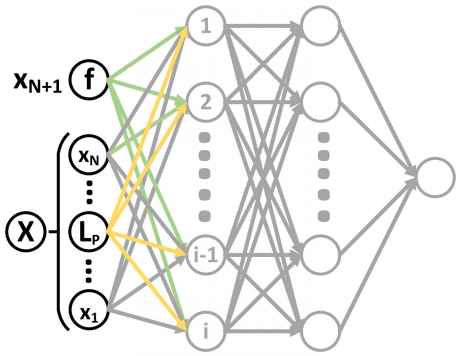

FIGURE 5. The DNN structure with conventional input feature.

Additionally, we can find a neural network structure represented by $g(x)$ in (7) that maps the given inputs $\vec{x}$ and L_e to the frequency response output. The neural network structure with the proposed electrical length, L_e, as an input feature is shown in Fig. 6.

$$y = g(\vec{x}, L_e) \tag{7}$$

In this structure, the proposed L_e is employed to substitute f into the conventional neural network. Moreover, the complexity of these models is identical, with the same hidden layers, number of neurons, and activation function for direct comparison. We show that $h(x)$ with independent f and L_p cannot obtain the same performance as $g(x)$ with the proposed L_e.

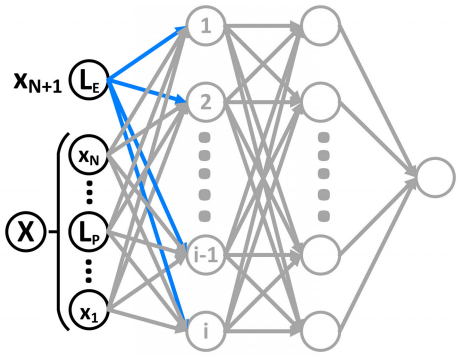

FIGURE 6. DNN structure with proposed L_e.

Function $h(x)$ described in Fig. 5 with independent L_p and f as inputs achieves the same performance as $g(x)$ represented in Fig. 6 if and only if the contribution of these inputs to $h(x)$ (determined by their weight) is equal to the one of L_e to $g(x)$. This condition is proven to be incorrect for all

L_p and f pairs. Hence, $h(x)$ will fail in broadband frequency response predictions with channel length variation.

Let's define $w^i_{j,k}$ and $u^i_{j,k}$ as the weight that connects the output of neuron j in layer (i-1) to input neuron k in layer i in $h(x)$ and $g(x)$, respectively. For each neuron k in the first hidden layer, the condition is described in (8).

$$w^1_{x_{N+1},k} \times f + w^1_{x_{N+2},k} \times L_p = u^1_{x_{N+1},k} \times L_e$$
$$f = \frac{w^1_{x_{N+2},k}}{u^1_{x_{N+1},k} - \frac{w^1_{x_{N+1},k}}{L_p}} \tag{8}$$

Neural network weights are deterministic or specific numbers selected after the training process. Hence, the weights used in (8) function as constant parameters to express the requirement for L_p and f. Consequently, given a specific channel length value, we can find a unique corresponding frequency value that satisfies the contribution condition in (8). In other words, if we change the channel length, the DNN model must find new weights that can preserve the contribution of L_e to achieve the same performance. This contradicts the property of deterministic weights in the DNN structure. Therefore, $h(x)$ cannot achieve the same performance as the proposed L_e in $g(x)$ regardless of the complex DNN structure.

In conclusion, the DNN model with existing techniques such as activation function or structure modification cannot conserve the properties caused by the electrical length in transmission line theory. If the channel length of the structure changes, the model will fail in its prediction without the information from the proposed L_e.

B. SNAKE ACTIVATION FUNCTION

In most cases, the strong performance of the DNN model is due to the proper activation function and structural complexity. While structural complexity can quickly increase by adding an increased number of hidden layers and neurons, choosing an appropriate activation function is critical in specific issues. For example, binary and multi-class classification problems typically use sigmoid, softmax, or tanh as activation functions, and the ReLU and its variants, such as parametric ReLU [41] (PReLU) are widely employed in CNN for image detection applications. The standard activation function is insufficient for modeling the periodic behavior, particularly in high-accuracy applications like frequency response prediction. In addition to the minimum overall error, the prediction curve must also match the simulation data in this scenario. In theory, the electrical length (L_e) significantly affects the local minimum and maximum. From (1), (2) and (5), the frequency response magnitude can be rewritten as (9).

$$|V_s(L_p)| = -\alpha L_p + \log(\cos(kL_e)) + C_0 \tag{9}$$

It is noticeable that term C_0 in (9) is constant because the magnitude of the excited signal and reflection coefficient are invariant in the broadband frequency. Utilizing Euler's formula and Taylor series, we can express the terms $\log(\cos(k \times L_e))$ in (9) by the sum of several periodical functions, as shown in (10).

$$\log(\cos(x)) = \sum_{m=1}^{\infty}(-1)^{m-1}\frac{\cos(2mx)}{m} - \log(2) \tag{10}$$

Notably, we are attempting to find an appropriate activation function for DNN to model the frequency response. Thus, a complicated activation function comprising high-order harmonic components is unnecessary to avoid the slow learning process because model accuracy can be addressed by thousands of weights in the DNN structure. As a result, in this case, first-order representation of the signal periodical properties is sufficient, as shown in (11).

$$|V_s(L_p)| \approxeq -\alpha L_p + \cos(2kL_e) + C_0 + \log(2) \tag{11}$$

The approximation formula of the signal in (11) introduces a linear reduction αL_p with a periodic component $(\sin(kL_e))^2$ that depends on L_e. Thus, it is critical to leverage appropriate activation functions to capture this property that typical activation functions cannot do. Hence, we leverage the snake activation function introduced in [34] to predict the frequency response with periodic properties. The snake activation function is described using the following equation:

$$y = -(x + \frac{1}{\theta}(\sin\theta x)^2) \tag{12}$$

where:

θ : parameter to control the frequency of periodic signal

C. RESONANT FEATURE FOR RESONANCE MODELING

In previous works, the channel length is fixed, so the resonant frequency is invariant in all training and test data sets. Nonetheless, the model often ignores errors from these spots because they are outliers from the machine learning perspective due to their irregular appearances. Consequently, the model cannot detect or reveal the potential resonant frequency in the presence of channel length variation, failing to deliver essential information to the designer.

In theory, for each integer N of L_e, we can find one resonant frequency for each given the physical channel length. Even if the frequency step is small, we can only observe the gradual magnitude decrease around the estimated resonant frequency. Hence, it is challenging to precisely identify resonant frequency. However, the rate of loss around the resonant frequency is remarkably greater than in other regions. Therefore, we define a resonance range based on this fact instead of trying to detect exactly one resonant frequency on the basis of the two main advantages.

First, using numerous frequencies within the predefined range instead of one frequency relax the learning process. Second, each frequency point in the resonance range is assigned a weight that describes the drop in loss to the next frequency point. When the frequency is near the

resonant frequency, the weight value will be larger, which means that the loss drop rate is high. Hence, the weight of resonant frequency corresponding to the local minimum in the resonance range will be 1. For other points in the resonance range, the weights gradually decreased to 0.5, indicating a decrease in the loss drop rate if the operating frequency is far from the resonant ones. Finally, a frequency with a location outside of the resonance range is assigned a weight of 0. The resonant feature extraction process is shown in Fig. 7. The modeling process can be summarized as follows.

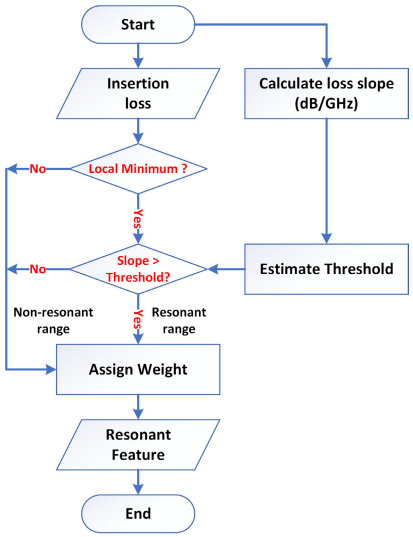

FIGURE 7. Summary of the resonant feature extraction flow.

Step 1: For each k^{th} training structure, calculate the slope (SL) of channel loss (IL) at each frequency point f_k^i as follows.

$$SL_k^i = \left| \frac{IL_k^i - IL_k^{i+1}}{f_k^i - f_k^{i+1}} \right| \tag{13}$$

Step 2: For each k^{th} training structure containing (N+1) frequency points, calculate a threshold using (14) to determine the resonant range using the loss slope in Step 1. The ideal loss in a channel is generally inversely proportional to the frequency across the broadband. Thus, we exploit the RMS value as a general guide for identifying the boundary between resonant and non-resonant areas. However, the strength factor ($C \geq 2$) can be adjusted to regulate the desired areas of resonance for a smooth curve.

$$SL_k^{thres} = \sqrt[C]{\frac{\sum_{i=1}^{N} (SL_k^i)^C}{N}} \tag{14}$$

Step 3: In each k^{th} training structure, determine the possible resonant frequencies by collecting the local minimum insertion loss.

Step 4: Refine the resonant range R using the loss slope in step 1 and threshold estimation in Step 2 for each possible resonant frequency f_k^i determined in Step 3.

$$f_k^i \begin{cases} \in R, & \text{if } SL_k^i \geq SL_k^{thres} \\ \notin R, & \text{if } SL_k^i < SL_k^{thres} \end{cases} \tag{15}$$

Step 5: In each resonant range R, we have the smallest loss l_k^low and highest loss l_k^high on the left side and right side of the resonant frequency, respectively. The weight w_k^i of each frequency point in this range is assigned as shown in (16).

$$\begin{aligned} \Delta L_l &= \left| IL_k^{res} - IL_k^{low} \right| \\ \Delta L_r &= \left| IL_k^{res} - IL_k^{high} \right| \\ w_k^i &= \begin{cases} \dfrac{IL_k^i - IL_k^{res}}{\Delta L_l}, & \text{if } f_k^i < f_{res} \\ \dfrac{IL_k^i - IL_k^{res}}{\Delta L_r}, & \text{if } f_k^i > f_{res} \\ 1, & \text{if } f_k^i = f_{res} \\ 0, & \text{if } f_k^i \notin R \end{cases} \end{aligned} \tag{16}$$

Fig. 8 shows an instance of resonant feature modeling. The red curve depicts a sample of insertion loss while the blue curve illustrates the assigned weight. At points of integer L_e, the resonant feature leads to the most significant loss drop, hence, the weight is 1. The weight decreases gradually to 0.5 around the resonant point. Otherwise it's 0.

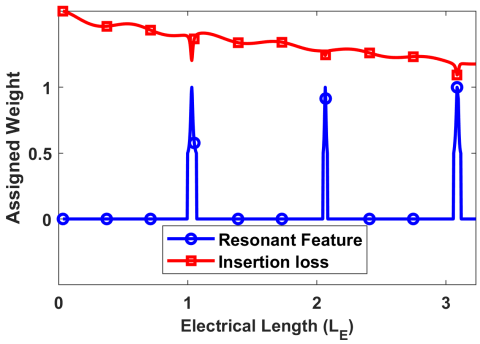

FIGURE 8. Example of resonant modeling.

D. OVERALL DESIGN FLOW

In this section, the overall design flow is summarized in Fig. 9. Our proposed methodology includes three stages, data generation, data pre-processing, and model training. In the first stage, the design space is sampled from DoE following the Latin hypercube sampling (LHS) method [40] before being fed into the EM simulator to generate channel frequency response data. In the second stage, we extract the resonant feature from the frequency response data following the process introduced in the Section III-C. This feature incorporates frequency response data to assemble output features that comprehensively describe the characteristics of the channel in the frequency domain. In addition, we carry out the proposed L_e based on physical channel length (L_p),

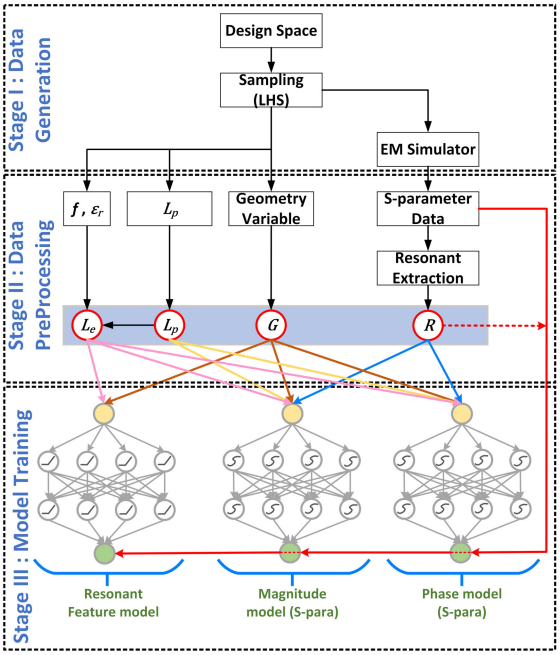

FIGURE 9. Summary of the proposed design flow.

frequency (f), and relative permittivity (ε_r). In the third stage, the surrogate model is comprised of three modules: resonant feature, magnitude, and phase modules. We can train these models simultaneously because all training input and output data are available. The resonant feature DNN-based model that predicts the resonant frequencies for channel frequency response takes the proposed electrical length L_e and geometry variables (represented by node G in Fig. 9) as the input vector due to the characteristics expressed in (4). Also, this model training output is the resonant feature extracted from frequency response training data in the second stage (represented by the node R in Fig. 9). We leverage the ReLU function as the activation function of hidden layers for this model. On the other hand, the resonant feature R electrical length L_e, physical length L_p, and geometry variables G are the input vector to train the magnitude and phase DNN-based models. The training outputs of these models are obtained from the frequency response training data, respectively. The proposed snake activation function is utilized as an activation function in the hidden layers in both these models. Also, all the DNN-based models in this work are trained using the Adam optimizer. Additionally, we spend only 80% of the provided training data to feed the DNN-based models for training purpose. Besides, the remaining 20% is reserved for continually testing models' performance, which can avoid overfitting and ensure generalization.

Next, the trained DNN-based models are validated by predicting the frequency response of the previously unseen validation data. In the prediction process, the resonant feature model first takes the input vector comprising L_e and G to predict the resonant feature. We also defined the vector of resonant feature predicted by this model as R, which is employed to represent the resonant feature extracted from frequency response data in the training process. Then, the phase and magnitude DNN-based models take an input vector containing L_e, L_p, G, and R to estimate a frequency response with predicted resonance frequencies for a given design structure.

For this regression problem, a variant of the mean squared error is applied in the learning process, as shown in (17). Positive terms α and β emphasize the significance of the high error in a prediction. The parameter β can be considered a threshold. If the prediction error is greater than the predefined value of β, the negative of term $(\beta - |y - \hat{y}|)$ increases its contribution to the overall loss. If the difference in prediction is less than β, the weight of this error will smaller.

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{(y_i - \hat{y}_i)^2}{1 + \exp^{\alpha(\beta - |y_i - \hat{y}_i|)}}\right) \tag{17}$$

where:

y_i : The real value
$\hat{y}_i$: The prediction value
N : Number of data points

We define the performance criteria to measure the DNN prediction performance to predict the S-parameter over a given frequency range. The fundamental metric for evaluating predictive accuracy is the residual or mean amplitude error (MAE). The mean amplitude percentage error (MAPE) was also computed to allow for a direct comparison to existing approaches. Because of the periodic features, the difference between the prediction and simulation for the phase is divided by 360°. The following equation describes the MAE, and MAPE.

$$\begin{aligned} MAE &= \frac{1}{D}\sum_{i=1}^{D}\left(\frac{1}{F_i}\sum_{j=1}^{F_i}\left|y_i^j - \hat{y}_i^j\right|\right) \\ MAPE &= \frac{1}{D}\sum_{i=1}^{D}\left[\frac{1}{F_i}\sum_{j=1}^{F_i}\left(\left|\frac{y_i^j - \hat{y}_i^j}{y_i^j}\right| 100\right)\right] \end{aligned} \tag{18}$$

where:

y_i^j : The simulation data at the j^{th} point in the i^{th} structure
$\hat{y}_i^j$: The prediction data at the j^{th} point in the i^{th} structure
F_i : Number of frequency points in the i^{th} structure
D : Number of structures

IV. NUMERICAL RESULTS AND ANALYSIS

In this section, we will examine the performance of various high-speed channel structures, including microstrip and serpentine designs. The reference S-parameter data of the structure, which consists of both magnitude and phase information, is exported from the ANSYS HFSS tool [39] integrated into the ANSYS Electronics Desktop.

A. MICROSTRIP LINE DESIGN

The microstrip line structure shown in Fig. 10 is employed to evaluate the performance of the proposed method. In addition to the channel width and space variation, the channel length is also varied over a wide range to verify the effectiveness of the proposed method.

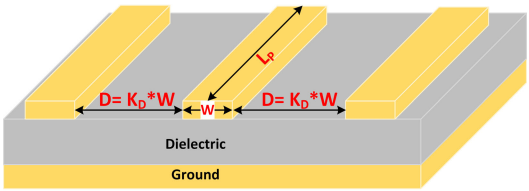

FIGURE 10. Microstrip line design example.

This microstrip line structure includes three parallel channels to measure the crosstalk effect. The dielectric material in the design is FR4 with a relative permittivity of 4.4. The channel width (W) ranges from 20μm to 70μm, and K_D is the ratio of channel space to channel width. This ratio is assigned a value between 1 and 3 which is typical in interconnect design to conserve area while minimizing crosstalk. Additionally, the channel length varies in a wide range of 20*mm* to 60*mm*. The frequency range is from 0.1 *GHz* to 10*GHz* with a step size of 0.01 *GHz*. Over this design space, 45 training data sets and 36 validation data sets covering the frequency range have been produced. The training data contains 44595 data points. Each data point consists of one magnitude value and one phase value corresponding to each other. We use 80% of the training data to train the DNN-based models while the remaining 20% is utilized for testing purposes. Then, the trained models are validated using the validation data. Table 1 shows a summary of the training and validation data.

TABLE 1. Training and validation data of microstrip line structures.

Parameters	*Min*	*Max*
Frequency (GHz)	*0.1*	*10*
W (μm)	*20*	*70*
K_D	*1*	*3*
L_p **(mm)**	*20*	*60*
No. of training datasets	*45*	
No. of validation datasets	*36*	

The proposed LD-DNN in this work shows an overall error of less than 4% for all four S-parameter magnitudes over the broadband frequency, as summarized in Table 2. Additionally, the frequency response of the magnitude and phase over the broadband frequency is depicted in Fig. 11 and Fig. 12. We use a DNN model with three hidden layers to predict the resonant frequencies. DNN models with five hidden layers are also employed to model the magnitude of the frequency response. Similarly, DNN models with five hidden layers represent the phase of frequency response. The proposed LD-DNN exhibits remarkable efficiency, enabling rapid prediction of the frequency response for any arbitrary

TABLE 2. The S-parameter prediction performance for microstrip line structures.

		Training Error		Validation Error	
		MAPE(%)	MAE	MAPE(%)	MAE
Mag.(dB)	S_{11}	1.49	0.3	3.31	0.76
	S_{21}	1.0	0.022	1.54	0.05
	S_{31}	1.12	0.35	2.66	0.78
	S_{41}	0.83	0.15	0.90	0.15
Phase(°)	S_{11}	0.96	3.3	2.75	9.36
	S_{21}	0.21	0.73	0.29	0.98
	S_{31}	1.79	6.12	2.12	7.24
	S_{41}	0.31	1.08	0.85	2.90

TABLE 3. Run time comparison of microstrip line structures.

Design Length (mm)	*LD-DNN*	*HFSS Simulator*
23	*1.24s*	*681.5s*
37	*1.35s*	*5670.9s*
52	*1.30s*	*29061.9s*

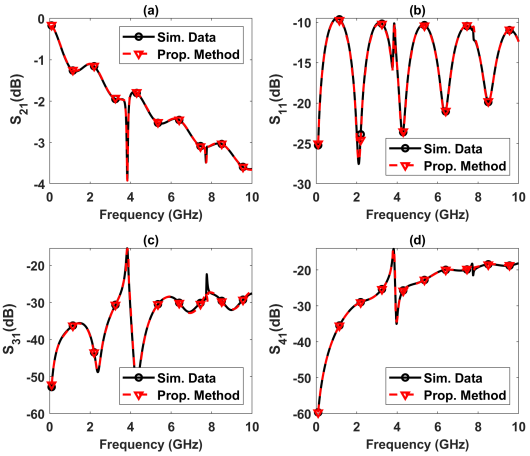

FIGURE 11. S-parameter magnitude prediction of microstrip line structures. (a) S_{21}. (b) S_{11}. (c) S_{31}. (d) S_{41}.

structure within 1.4 seconds, as shown in Table 3. Furthermore, the proposed LD-DNN offers a notable advantage in that its runtime for generating the frequency response remains independent of the channel length, distinguishing it from the HFSS simulator. Consequently, the steady accuracy and speed of this LD-DNN, regardless of the channel's size, makes it a valuable and practical tool for high-speed channel modeling.

While the high-speed channel is one of the most time-consuming designs, there have not been many efforts in utilizing machine learning to optimize this procedure. So we compared our method with those developed for antenna applications [19], [20], [21] and for high-speed channels in [26] and [27]. However, the comparison results reveal that methods developed for antenna applications are inappropriate for a high-speed channel. The technique proposed in [26] predicts the frequency response by extrapolating the RLGC components predicted by DNN models. Because the authors claimed the prediction inaccuracy of the proposed approach

TABLE 4. Performance comparison with previous works for microstrip line structures.

	Performance Evaluation				DNNs Structure		
	Train MAPE(%)	Valid. MAPE(%)	Reflection Effect	Resonance Effect	Hidden Layers	Hidden Neurons	Activation Function
Method *[19]*	12.0	11.1	x	x	6	40-80-120-120-80-40	Sigmoid
Method *[20]*	10.0	10.9	x	x	6	40-80-120-120-80-40	Sigmoid
Method *[21]*	10.4	10.5	x	x	8	100-100-100-100-100-100-100-100	Sigmoid, SReLU
Method *[26]*	-	13.8	✓	x	-	-	-
Method *[27]*	7.6	8.8	x	x	6	40-80-120-120-80-40	PReLU
Prop. LD-DNN	1.0	1.5	✓	✓	5	32-64-96-64-32	Snake

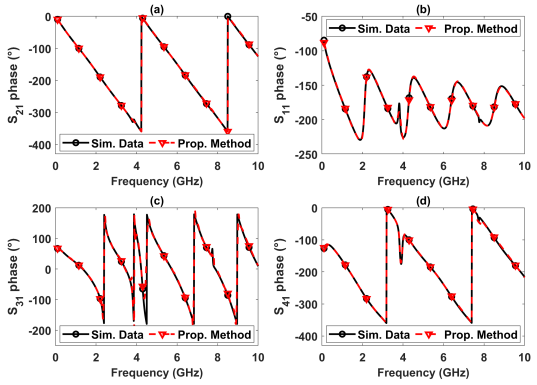

FIGURE 12. S-parameter phase prediction of microstrip line structures. (a) S_{21}. (b) S_{11}. (c) S_{31}. (d) S_{41}.

was less than 5%, we also employ the RLGC components directly extracted from the EM simulator with a variation of 1% to extrapolate the frequency response for this comparison. For a fair comparison with the proposed approach, the techniques in [19], [20], [21], and [27] are all implemented, and then they are trained and tested on the same data set in this work. For accuracy criteria, we employ the MAPE introduced in (18). More importantly, the multiple reflections and resonant frequency prediction are also examined.

The summary of the performance and structure comparison between the different DNN-based approaches are shown in Table 4. In this experiment, the methods published in [19], [20] show a large validation MAPE of 11.1% and 10.9% in insertion loss, respectively. The method designed for high-speed channels presented in [27] produces an error of approximately 9% and the technique published in [26] which is based on the RLGC components from the EM simulator (Q2D) shows an exceptionally high error of 13.8%. More importantly, neither proposed techniques for high-dimensional structure applications nor high-speed channels can effectively capture the multiple reflection effect and resonant frequencies in a broad frequency band, as illustrated in Fig. 13. The high error observed in previous techniques can be attributed to two primary reasons. Firstly, these approaches cannot adequately account for the multiple reflection effect occurring along the channel, mainly due to the use of inappropriate activation functions. Secondly, these methods face challenges in capturing the resonant frequencies because they do not have access to accurate information provided by L_e. For instance, specific resonances may manifest around 4 and 8 GHz; these are only captured by the proposed LP-DNN as illustrated in Fig. 13. In contrast to the proposed method in which a smaller model size is employed, these other approaches utilize significantly larger models with more hidden layers and hidden neurons. Despite this advantage in model size, their prediction inaccuracies are still beyond 5%. Hence, our proposed LP-DNN outperforms existing approaches in terms of the numerical performance with a 1.5% error. Consequently, the proposed LD-DNN can be utilized without the need for additional training in the event of varying the channel length. Conversely, existing approaches must be started from the beginning, which involves generating new training data and conducting the fitting process once more.

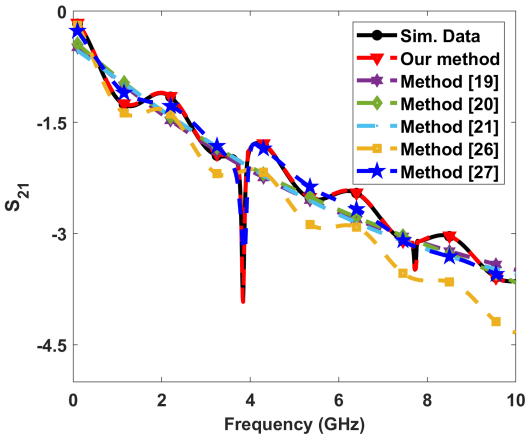

FIGURE 13. Insertion loss prediction comparison for microstrip line structures.

B. SERPENTINE INTERCONNECT DESIGN

In this section, we evaluate our method on a serpentine interconnect structure, as shown in Fig. 14. Serpentine structures can control the coupling inductance and capacitance to minimize crosstalk in high-speed channels by adjusting the number of projecting parts [37]. In our test case, the number of projecting parts, or number of blocks, ranges from 1 to 50. This design employs FR4 with a relative permittivity of 4.4 for the dielectric material. For the cross-section,

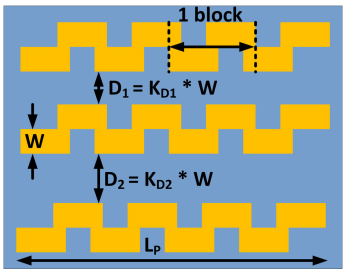

FIGURE 14. Serpentine interconnect design.

the channel width (W) range is from 20μm to 100μm. K_{D1} ranges from 0.5 to 3.5 and K_{D2} ranges from 0.5 to 5.5 in order to represent the distance between central channels and others. The value of K_{D2} is always greater than K_{D1} in all structures. The channel length ranges from 10*mm* to 30*mm*, and the frequency ranges from 0.1 *GHz* to 10*GHz* with a 0.01 GHz step size. Over this design space, 90 training and 72 validation data sets have been produced and data information is summarized in Table 5. In this example, we employ 80% training data equivalent to 71352 values of magnitude and phase corresponding to each other to train DNN-based models. Besides, the remaining 20% of training data is used for testing. Then, validation data is utilized to validate the trained models.

TABLE 5. Training and validation data of serpentine structures.

Parameters	*Min*	*Max*
Frequency (GHz)	0.1	10
W (μm)	20	100
K_{D1}	0.5	3.5
K_{D2}	0.5	5.5
L_p (mm)	10	30
Block	1	50
No. of training datasets	90	
No. of validation datasets	72	

The model results are compared with 3D HFSS results and show an overall error of approximately 3% for the prediction of all frequency response components. The magnitude and phase responses are compared in Fig. 15 and Fig. 16.

The overall accuracy and time-efficiency are summarized in Tables 6 and 7. The proposed LD-DNN can predict the frequency response in approximately 1.4 seconds. This represents a substantial benefit when compared with the HFSS simulator, as the simulation time in HFSS increases exponentially with the channel length.

The insertion loss prediction differences between the different DNN-based approaches for serpentine structures are shown in Fig. 17, and a summary of the performance comparison is shown in Table 8. The high speed channel modeling method proposed in [26], which relies on RLGC components from a 2D simulator, is inapplicable to this structure due to the 3-D geometry. The inaccuracies of the previous methods originated from inaccuracies in modeling multiple reflection effects and resonant frequencies in a broad frequency spectrum. In this comparison, it is worth noting

TABLE 6. The S-parameter prediction performance of serpentine structures.

		Training Error		Validation Error	
		MAPE(%)	MAE	MAPE(%)	MAE
Mag.(dB)	S_{11}	0.46	0.07	3.11	0.56
	S_{21}	0.65	0.006	2.14	0.03
	S_{31}	0.54	0.21	1.34	0.48
	S_{41}	0.31	0.11	0.90	0.31
Phase(°)	S_{11}	0.22	0.76	2.53	8.84
	S_{21}	0.33	1.17	0.61	2.12
	S_{31}	0.39	1.37	1.24	4.33
	S_{41}	0.47	1.65	1.46	5.11

TABLE 7. Runtime comparison of serpentine structures.

Design Length (mm)	*LD-DNN*	*HFSS Simulator*
12	*1.32s*	*705.0s*
18	*1.36s*	*1002.1s*
27	*1.38s*	*2077.83s*

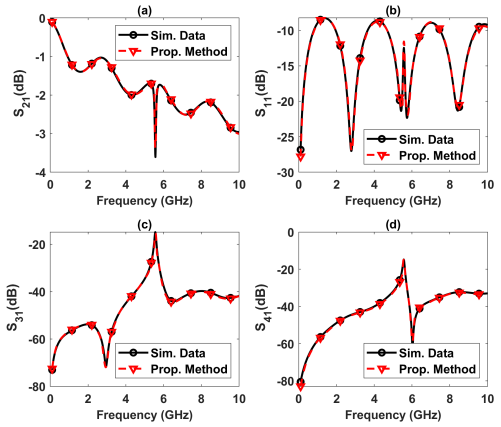

FIGURE 15. S-parameter magnitude prediction of serpentine lines in the broadband frequency. (a) S_{21}. (b) S_{11}. (c) S_{31}. (d) S_{41}.

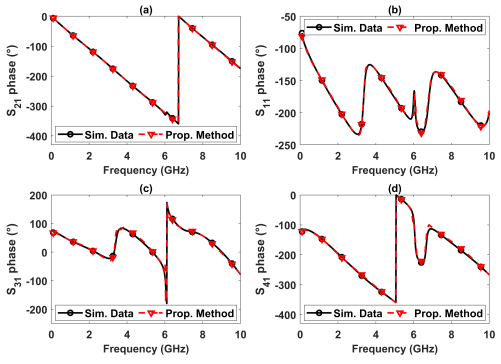

FIGURE 16. S-parameter phase prediction of serpentine lines in the broadband frequency. (a) S_{21}. (b) S_{11}. (c) S_{31}. (d) S_{41}.

that the DNNs proposed in [19], [20], and [21] have been implemented with a significantly higher number of hidden neurons, and yet they demonstrate similar performance for both the training and validation datasets. There can be some limitation in using these techniques in high-speed channel

TABLE 8. Performance comparison with previous works in serpentine interconnect structure.

	Performance Evaluation				DNNs Structure		
	Train MAPE(%)	Valid. MAPE(%)	Reflection Effect	Resonance Effect	Hidden Layers	Hidden Neurons	Activation Function
Method *[19]*	12.7	13.5	x	x	6	30-60-90-90-60-30	Sigmoid
Method *[20]*	25.5	25.9	x	x	6	30-60-90-90-60-30	Sigmoid
Method *[21]*	11.1	13.0	x	x	6	100-100-100-100-100-100	Sigmoid, SReLU
Method *[27]*	3.5	12.7	x	x	5	24-36-48-36-24	PReLU
Prop. LD-DNN	0.65	2.1	✓	✓	5	24-36-48-36-24	Snake

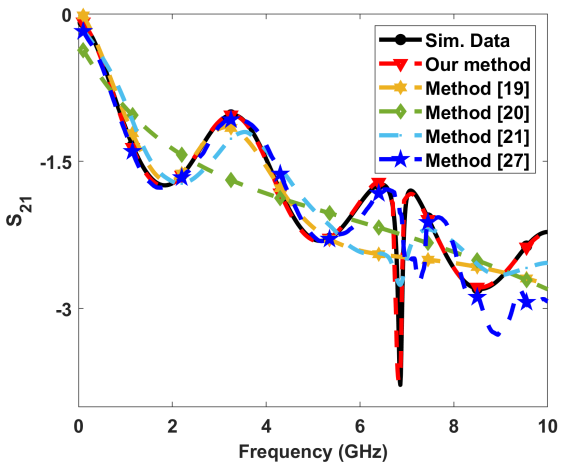

FIGURE 17. Insertion loss prediction comparison in serpentine structures.

modeling, as increasing the model size further could lead to over-fitting.

V. CONCLUSION

In this article, we present an LD-DNN-based modeling methodology that accurately and efficiently predicts the length-dependent high-speed channel frequency response. The length dependent characteristics can only be captured by 3D EM simulations which are very time consuming. Our method can generate a surrogate model that captures the frequency response of the high-speed channel as the length changes. We propose employing the electrical length as an input feature and adopting the snake activation function in DNN to capture the multiple reflections caused by the impedance mismatch. Additionally, we propose the resonant feature to model the resonance effects caused by crosstalk from neighboring signal lines. Compared with existing methods, the proposed technique uses fewer training parameters in DNN while significantly improving the accuracy of channel loss prediction. More importantly, the proposed method can accurately predict the multiple reflection effects by capturing periodic components in the frequency response if there are impedance mismatches. The proposed method was tested on microstrip and serpentine channel structures, and the results show significant improvement in terms of the insertion loss prediction as compared to previously published results. With this model, the optimum channel design can be found efficiently, and if it is combined with high-speed transceiver circuits, overall system design optimization can be achieved at the system level.

REFERENCES

[1] Q.-J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. Norwood, MA, USA: Artech House, 2000.
[2] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 420–435, Jan. 2004.
[3] H. Kabir, L. Zhang, M. Yu, P. H. Aaen, J. Wood, and Q.-J. Zhang, "Smart modeling of microwave devices," *IEEE Microw. Mag.*, vol. 11, no. 3, pp. 105–118, May 2010.
[4] D. E. Root, "Future device modeling trends," *IEEE Microw. Mag.*, vol. 13, no. 7, pp. 45–59, Nov. 2012.
[5] G. Huang, J. Hu, Y. He, J. Liu, M. Ma, Z. Shen, J. Wu, Y. Xu, H. Zhang, K. Zhong, X. Ning, Y. Ma, H. Yang, B. Yu, H. Yang, and Y. Wang, "Machine learning for electronic design automation: A survey," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 26, no. 5, pp. 1–46, Sep. 2021.
[6] H. Wang, J. Yang, H.-S. Lee, and S. Han, "Learning to design circuits," in *Proc. NeurIPS Workshop*, 2018. [Online]. Available: https://arxiv.org/abs/1812.02734
[7] H. Wang, K. Wang, J. Yang, L. Shen, N. Sun, H.-S. Lee, and S. Han, "GCN-RL circuit designer: Transferable transistor sizing with graph neural networks and reinforcement learning," in *Proc. 57th ACM/IEEE Design Autom. Conf. (DAC)*, Jul. 2020, pp. 1–6.
[8] K. Settaluri, A. Haj-Ali, Q. Huang, K. Hakhamaneshi, and B. Nikolic, "AutoCkt: Deep reinforcement learning of analog circuit designs," in *Proc. Des. Autom. Test Eur.*, Mar. 2020, pp. 490–495.
[9] H. M. Torun, M. Swaminathan, A. Kavungal Davis, and M. L. F. Bellaredj, "A global Bayesian optimization algorithm and its application to integrated system design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 792–802, Apr. 2018.
[10] H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2128–2142, Jun. 2019.
[11] H. M. Torun, M. Larbi, and M. Swaminathan, "A Bayesian framework for optimizing interconnects in high-speed channels," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2018, pp. 1–4.
[12] N. Ambasana, D. Gope, B. Mutnury, and G. Anand, "Automated frequency selection for machine learning based EH/EW prediction from S-parameters," in *Proc. 24th IEEE Elect. Perform. Electron. Packag. Syst.*, Oct. 2015, pp. 53–56.
[13] N. Ambasana, G. Anand, B. Mutnury, and D. Gope, "Eye height/width prediction from *S*-parameters using learning-based models," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 6, no. 6, pp. 873–885, Jun. 2016.
[14] N. Ambasana, G. Anand, D. Gope, and B. Mutnury, "S-parameter and frequency identification method for ANN-based eye-height/width prediction," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 7, no. 5, pp. 698–709, May 2017.
[15] K. S. Ooi, C. L. Kong, C. H. Goay, N. S. Ahmad, and P. Goh, "Crosstalk modeling in high-speed transmission lines by multilayer perceptron neural networks," *Neural Comput. Appl.*, vol. 32, no. 11, pp. 7311–7320, Jun. 2020.
[16] F. Feng, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60–77, Jan. 2016.

[17] F. Feng, V.-M.-R. Gongal-Reddy, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of microwave components using adjoint neural networks and pole-residue transfer functions with EM sensitivity analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 1955–1975, Jun. 2017.
[18] J. Zhang, F. Feng, W. Zhang, J. Jin, J. Ma, and Q.-J. Zhang, "A novel training approach for parametric modeling of microwave passive components using Padé via Lanczos and EM sensitivities," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 6, pp. 2215–2233, Jun. 2020.
[19] J. Dong, W. Qin, and J. Mo, "Low-cost multi-objective optimization of multiparameter antenna structures based on the l_1 optimization BPNN surrogate model," *Electronics*, vol. 8, no. 8, p. 839, Jul. 2019.
[20] J. Jin, F. Feng, J. Zhang, S. Yan, W. Na, and Q. Zhang, "A novel deep neural network topology for parametric modeling of passive microwave components," *IEEE Access*, vol. 8, pp. 82273–82285, 2020.
[21] J. Jin, C. Zhang, F. Feng, W. Na, J. Ma, and Q.-J. Zhang, "Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 4140–4155, Oct. 2019.
[22] J. Sun, Y. Hu, H. Fang, and Z. Wang, "Accurate S parameter prediction of L-shaped probe-fed patch antenna with an improved artificial bee colony algorithm based on artificial neural network," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 31, no. 9, Sep. 2021, Art. no. e22783.
[23] H. M. Torun, A. C. Durgun, K. Aygün, and M. Swaminathan, "Causal and passive parameterization of S-parameters using neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 10, pp. 4290–4304, Oct. 2020.
[24] H. M. Torun, A. C. Durgun, K. Aygün, and M. Swaminathan, "Enforcing causality and passivity of neural network models of broadband S-parameters," in *Proc. IEEE 28th Conf. Electr. Perform. Electron. Packag. Syst. (EPEPS)*, Oct. 2019, pp. 1–3.
[25] H. M. Torun, H. Yu, N. Dasari, V. C. K. Chekuri, A. Singh, J. Kim, S. K. Lim, S. Mukhopadhyay, and M. Swaminathan, "A spectral convolutional net for co-optimization of integrated voltage regulators and embedded inductors," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2019, pp. 1–8.
[26] H. Kim, C. Sui, K. Cai, B. Sen, and J. Fan, "Fast and precise high-speed channel modeling and optimization technique based on machine learning," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 6, pp. 2049–2052, Dec. 2018.
[27] H. Khac Le and S. Kim, "Machine learning based energy-efficient design approach for interconnects in circuits and systems," *Appl. Sci.*, vol. 11, no. 3, p. 915, Jan. 2021.
[28] B. Pu, H. Kim, X.-D. Cai, B. Sen, C. Sui, and J. Fan, "Training set optimization in an artificial neural network constructed for high bandwidth interconnects design," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 6, pp. 2955–2964, Jun. 2022.
[29] W. C. Na and Q. J. Zhang, "Automated knowledge-based neural network modeling for microwave applications," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 7, pp. 499–501, Jul. 2014.
[30] W. Na and Q.-J. Zhang, "Unified automated knowledge-based neural network modeling for microwave devices," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Ottawa, ON, Canada, Aug. 2015, pp. 1–3.
[31] W. Na, F. Feng, C. Zhang, and Q.-J. Zhang, "A unified automated parametric modeling algorithm using knowledge-based neural network and l_1 optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 3, pp. 729–745, Mar. 2017.
[32] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. Int. Conf. Mach. Learn.*, 2003, pp. 1–8.
[33] K. S. Balagani and V. V. Phoha, "On the feature selection criterion based on an approximation of multidimensional mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1342–1343, Jul. 2010.
[34] L. Ziyin, T. Hartwig, and M. Ueda, "Neural networks fail to learn periodic functions and how to fix it," in *Proc. 34th NeurIPS*, 2020, pp. 1583–1594.
[35] S. B. Smith and M. Rengarajan, "Overcoming the Si challenges in designing 25–40 Gb/s backplane channels," in *Proc. DesignCon*, May 2014, pp. 953–978.
[36] J. Ferry and G. Biddle, "Optimizing symmetry in open field designs," in *Proc. DesignCon*, 2015. [Online]. Available: https://suddendocs.samtec.com/notesandwhitepapers/8-we6paper_optimizingsymmetry inopenfield.pdf
[37] K. Lee, H.-K. Jung, H.-J. Chi, H.-J. Kwon, J.-Y. Sim, and H.-J. Park, "Serpentine microstrip lines with zero far-end crosstalk for parallel high-speed DRAM interfaces," *IEEE Trans. Adv. Packag.*, vol. 33, no. 2, pp. 552–558, May 2010.
[38] D. M. Pozar, *Microwave Engineering*, 4th ed. New York, NY, USA: Wiley, 2012.
[39] *ANSYS Electronics Desktop*, ANSYS, Canonsburg, PA, USA, 2019. [Online]. Available: https://www.ansys.com/products/electronics
[40] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, p. 239, May 1979.
[41] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.

HUNG KHAC LE (Student Member, IEEE) received the joint B.S. degree in electronic and communication engineering from the University of Science and Technology, and the University of Da Nang, in 2016. He is currently pursuing the M.S. and Ph.D. degrees with the Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, South Korea.

His current research interests include power management IC design, high speed I/O design, signal integrity, and electromagnetic compatibility.

SOYOUNG KIM (Senior Member, IEEE) received the B.S. degree in electrical engineering from Seoul National University, Seoul, Republic of Korea, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1999 and 2004, respectively.

From 2004 to 2008, she was with Intel Corporation, Santa Clara, CA, where she worked on parasitic extraction and simulation of on-chip interconnects. From 2008 to 2009, she was with Cadence Design Systems, San Jose, CA, where she worked on developing IC power analysis tools. She is currently a Professor with the Department of Semiconductor Systems Engineering, College of Information and Communication Engineering, Sungkyunkwan University, Suwon, Republic of Korea. Her research interests include VLSI computer-aided design, signal integrity, power integrity, and electromagnetic interference in electronic systems.