This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2019.2948527, IEEE Transactions on Circuits and Systems II: Express Briefs

Semi-supervised learning based on Hybrid Neural Network for the Signal Integrity Analysis

Siyu Chen, *Student Member, IEEE,* Jienan Chen*, *Member, IEEE* Tingrui Zhang, *Student Member, IEEE,* Shuwu Wei, *Student Member, IEEE,*

***Abstract*—The signal integrity analysis of high-speed circuit channels becomes a challenging task, with the development of integrated circuit technology. To solve this problem, we proposed a fast-training semi-supervised learning method based on hybrid neural network (HNN) to predict the eye-diagram metrics. Compared with the existing methods, the proposed method only requires a small amount of training data with labels, the proposed method can automatically generate the labels for the unlabeled data with a small amount of labeled data with HNN based semi-supervised learning. To this end, the proposed method can save a great amount of time, which will be a more realistic solution for the practical application. Compared with existing machine learning-based methods, the proposed method requires 50% less labeled data for training with 32.29% and 20.73% accuracy improving on deep neural network (DNN) and co-training-style semi-supervised regression (COREG) methods, receptively.**

***Index Terms*—eye-diagram predict, semi-supervised learning, signal integrity analysis, hybrid neural network (HNN)**

I. Introduction

With the persistent upgrading of electronic products, electronic components are developing toward high speed and large scale [1]. Signal integrity analysis is vital to the evaluation of a digital product. The total jitter (TJ), signal to noise ratio (SNR) and the performance of anti-jamming are reflected by signal integrity [2].

Traditional simulation methods of inspecting signal integrity can be expensive and time-consuming [3]. Research on utilizing circuit physical properties to improve signal integrity analysis accuracy in the silicon interposer has been proposed [4]. Besides, there are efforts enhancing the ability of electromagnetic solvers by applying step response of the system. Electromagnetic solvers and circuit simulators are two major types of techniques to predict eye-diagram metrics. [5]. In digital electronics, a stream of binary values is represented by a voltage (or current) waveform. With the development and upgrade of electronic products, integrated circuits and high-speed signal propagation become keystones of modern electronics. There are increasing challenges in signal integrity analysis in high-speed systems. Electromagnetic solvers are affected by excessive factors such as component selection and board design etc. Circuit simulators are hard to consider adjacent switching signals and unrelated signals for integrity analysis [6].

Recently, machine learning methods are increasingly used to solve industry problems [7]–[11] *i.e.*, motor fault detection, pearl automatic sorting machine, dynamic system approximation and control etc, and it is a future trend to use machine learning to assist circuit automatic design(CAD). For signal integrity analysis, a machine learning-based method is proposed by Google with a regression model, which achieves about 90% accuracy [12]. However, since the labeled data can only be obtained from traditional signal integrity analysis methods, the acquisition of labeled data for training is time-consuming. This challenge will be more critical as the increasing of the prediction data dimension and requirement of accuracy.

To address this issue, some semi-supervised learning methods for applying both labeled data sets and unlabeled data sets are proposed in [13]–[16]. The idea of these methods is to label unlabeled data sets by cross-validation between multiple classifiers, which is unable to be applied in signal integrity analysis directly. In [17], the authors propose a single-view multi-regressor COREG method that can be applied to signal integrity analysis. However, these regressors are based on the same network structure, it exhibits poor performance on predicting eye-diagram metrics.

In this work, we propose a new semi-supervised learning for signal integrity analysis to predict the eye-diagram metrics, which is based on a hybrid neural network (HNN) to predict the eye-diagram metrics. The contributions of this work are summarized as follows:

1) Hybrid neural network-based automatically data labeling method. The core idea of HNN is the cross-validation of unlabeled data with two separate neural networks, which have different structures and trained with different distribution data. According to [18], if the inference results that output from two distinct structures and parameters distribution networks are with similar results, the probability of correct inference is much larger than a single network. Thus, we can employ the unlabeled data that passed cross-validation as the new training data for HNN.

2) Euclidean distance-based data validation method. We first employ an Euclidean distance calculator to validate the output of two neural networks and store these data into a validation data set. Then, we employ a stochastic sampling method to generate the training data set from the validation data set.

3) Significantly improvement in learning accuracy with less requirement for labeled data. The proposed method achieves the same accuracy as DNN [12], which can reduce the tagged

S. Chen, J. Chen*, T. Zhang, S. Wei was with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China e-mail: (Jesson.Chen@outlook.com).

The corresponding author is Jienan Chen

This work was supported by National Natural Science Foundation of China (61971107)

The paper was submitted on October 17, 2019.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2019.2948527, IEEE Transactions on Circuits and Systems II: Express Briefs

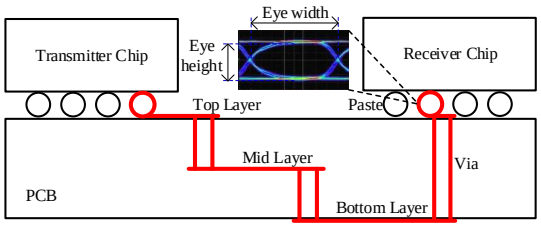

Fig. 1. The structure of chip-to-chip connections.

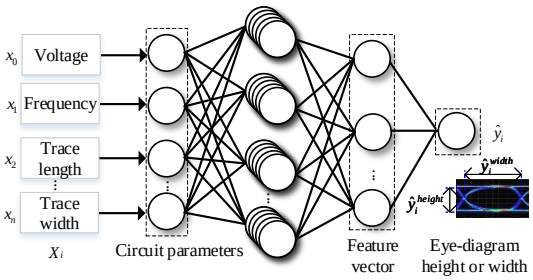

Fig. 2. The structure of DNN network.

data by 50% and improve eye-diagram prediction performance by 32.29%. Compared with the COREG [17] method, the eye-diagram height prediction performance is improved by 20.73%.

II. Hybrid Neural Network Method

A. Signal Integrity Analysis

As the frequency and complexity of high-performance system designs increase, various effects like crosstalk, reflection, and electromagnetic interferences can degrade the electrical signal to the point where errors occur and the system or device fails. Signal integrity analysis is the task of analyzing and mitigating these effects by measuring the quality of an electrical waveform. It is an essential activity at all levels of electronics packaging and assembly, especially for high-speed chip-to-chip system-level design, as shown in Fig. 1.

To evaluate signal integrity at the system level, most high-speed designs use eye-diagrams metrics. The transient waveform obtained by the circuit simulator is segmented and then superimposed to obtain an eye-diagram, and the signal integrity of the system is evaluated by the height y^{height} and width y^{width} of the eye-diagrams metrics.

B. Supervised Learning Based Signal Integrity Analysis

Supervised learning in [12] [19] is applied in signal integrity analysis of high-speed chip-to-chip systems. As shown in Fig. 2, the DNN takes n circuit parameters $X = \{x_0, x_1, \ldots, x_n\}$ of the high-speed channel as an input and outputs eye-diagram parameters (eye height y^{height} and eye width y^{width}) of the high-speed channel. The supervised learning required a large amount of labeled data for training the neural network, which is a time-consuming task for signal integrity analysis. For example, to generate 1K labeled training data would cost more than 200 hours by software Intel Integrated Channel Analysis Tool (ICAT). Hence, the acquiring

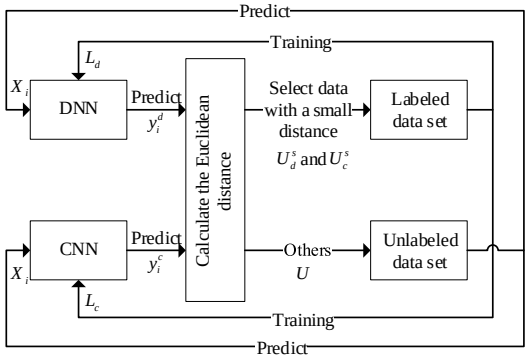

Fig. 3. The structure of HNN learning controller.

of labeled training data will be a bottleneck for the traditional supervised learning-based signal integrity analysis.

C. HNN Learning Model

In this section, we present a hybrid neural network (HNN) model based on semi-supervised learning to predict eye-diagram metrics of the high-speed chip-to-chip system. Semi-supervised can significantly reduce the amount of labeled data in training. It falls between unsupervised learning (unlabeled data, low accuracy) and supervised learning (labeled data, high accuracy). By applying unlabeled data in conjunction with a small amount of labeled data, semi-supervised learning can produce considerable improvement on learning accuracy over unsupervised learning. The HNN consists of two neural networks, a DNN and a convolutional neural network (CNN) [20] as shown in Fig. 3. We first apply a small amount of labeled data to train two networks until they return the output in reasonable accuracy. After that, we employ two trained networks to return the output of unlabeled data, and calculate the Euclidean distance between the output of the two networks. The output data from the two networks with small Euclidean distance are considered to be relatively more accurate, and being clustered to the valid data set and others are sent back to the unlabeled data pool. After obtaining a certain number of valid data, we can continue to train the two networks via data randomly sampled from the valid data set. Iteratively, more relatively accurate data is cluster with improved DNN and CNN. Hence, as iteration goes, the clustered data will be more accurate, which provides more labeled data with higher accuracy for training.

Before the first iteration, we sample two sub-sets L_d and L_c from a small amount of labeled data sets $L = \{(X_i, y_i)\}$ to pre-train the DNN and CNN. In the first iteration, the unlabeled data sets $U = \{(X_i)\}$ is fed to the pre-trained DNN and CNN to obtain the prediction label as y_i^d and y_i^c, respectively. To combine with the input unlabeled data, we can formulate two data sets as $U_d = \{(X_i, y_i^d)\}$ and $U_c = \{(X_i, y_i^c)\}$. Since the neural network is only trained with a small amount of data, the distribution of predicted results has a wide variance. The predicted data of two neural networks in overlap area with small Euclidean distance also have a large probability of laying in correct predicted label area. We can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2019.2948527, IEEE Transactions on Circuits and Systems II: Express Briefs

first calculate the Euclidean distance of the output results from the two networks as

$$dis = \left\{ \left| y_i^d - y_i^c \right|, i = K, K+1, ..., N \right\}. \quad (1)$$

We select data U_d^s and U_c^s with smaller Euclidean distance from the predicted data U_d and U_c. Then, we randomly sampled q data $U_d^{s\prime}$ and $U_c^{s\prime}$ from U_d^s and U_c^s at sampling rate γ, and remove them from U. In the next iteration, we add them into labeled data set separately, as

$$L_d = L_d \cup U_c^{s\prime},\ L_c = L_c \cup U_d^{s\prime}. \quad (2)$$

In the next step, DNN and CNN are trained with new data sets L_d and L_c. If the performance of the model is improved, the new data sets $U_d^{s\prime}$ and $U_c^{s\prime}$ are continued to be selected in the next iteration. Otherwise, the data added in the previous iteration is removed as

$$L_d = L_d \setminus U_c^{s\prime},\ L_c = L_c \setminus U_d^{s\prime} \quad (3)$$

Simultaneously, the new data sets $U_d^{s\prime}$ and $U_c^{s\prime}$ is formulated. The iteration process is performed until the convergence threshold is reached.

D. DNN and CNN Learning Model for HNN

As shown in Fig. 2, we improved the fully connected neural network structure of [12] for our system to learn eye-diagram performance metrics, with p hidden layers, input and output layers. The numbers of neurons for each layer are $H = \{H^{in}, H^1, ..., H^p, H^{out}\}$. The input data are circuit parameters of a high-speed channel system, formulated as vector X_i. The output data are eye height $\hat{y}^{height}$ and width $\hat{y}^{width}$. For instance, the output of all neurons in $l\ th$ hidden layer is:

$$Z^l = f_a^l(Z^{l-1}W^l + b^l). \quad (4)$$

In the formula (4), Z^{l-1} is the output from layer $l-1$, W^l is a weight matrix size of $H^{l-1} \times H^l$ and b^l is the bias matrix size of $1 \times H^l$. The parameters are presented as $\theta = \{(W^l, b^l), l = 0, 1, ..., q\}$, f_a is an activation function. In this design, we employ rectified linear unit (relu) function as the activation function

$$f_a(z) = \max(0, z). \quad (5)$$

The loss function of DNN is defined as:

$$J_d(\theta) = \sum_i^K (y_i - \hat{y}_i)^2, \quad (6)$$

As shown in Equation (7), DNN seeks to optimize regression $\hat{y}_i = f_\theta^d(X_i)$ by minimizing loss function $J_d(\theta)$.

$$\theta \leftarrow \arg\min_\theta \{J_d(\theta)\}. \quad (7)$$

For height and width, we use two independent neural networks to build the model as:

$$\hat{y}_i^{height} = f_{\theta_1}^d(X_i), \quad \hat{y}_i^{width} = f_{\theta_2}^d(X_i). \quad (8)$$

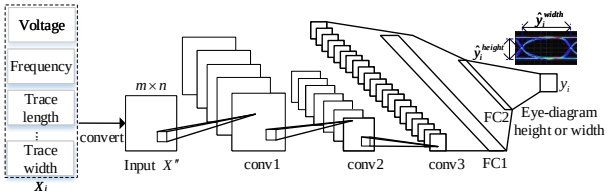

Fig. 4. The structure of CNN network.

As shown in Fig. 4, we proposed a deep convolutional neural network (CNN) to simulate eye-diagram metrics, which consists of three convolutional layers and two fully connected layers. Denote shape of convolution kernel for each layer as $C = \{C^1, C^2, C^3\}$, number of hidden nodes in full-connected hidden layers as $H = \{H^1, H^2\}$. To accommodate input characteristics of CNN,we treat X_i by:

$$X_i{}' = \left\{ x_{i,j}{}' = \left\lfloor \frac{\left(x_{i,j} - x_{:,j}^{\min}\right)}{\left(x_{:,j}^{\max} - x_{:,j}^{\min}\right)/m} \right\rfloor, j = 0, 1, ..., n \right\}, \quad (9)$$

where $x_{:,j}^{\max}$ and $x_{:,j}^{\min}$ represent the maximum and minimum value of all j-dimensional data in X_i, m is the number of segments for data. In this way, our data have been approximated to integral values between 0 and m. By using one-hot code to encode $X_i{}'$ we obtain matrices X_i^c with dimension $m \times n$, where each elements in column j is 0 except those at row $x_{i,j}{}'$ are 1. X_i^c is input into CNN, to output height $\hat{y}^{height}$ and width $\hat{y}^{width}$ of eye-diagram. We define the loss function of CNN as:

$$J_c(\theta) = \sum_i^K (y_i - \hat{y}_i)^2, \quad (10)$$

As shown in Equation (11), CNN seeks to optimize regression $\hat{y}_i = f_\theta^c(X_i)$ by minimizing loss function $J_c(\theta)$.

$$\theta \leftarrow \arg\min_\theta \{J_c(\theta)\}. \quad (11)$$

For height and width, we employ two independent neural networks to build the model as:

$$\hat{y}_i^{height} = f_{\theta_1}^c(X_i), \quad \hat{y}_i^{width} = f_{\theta_2}^c(X_i). \quad (12)$$

III. NUMERICAL EXAMPLES

A. Test Model

In this section, we perform a simulation on DDR4 model and set various parameters to evaluate the performance of this HNN method we proposed. Totally, $n = 45$ circuit parameters of high-speed channel are taken as input. Among them, $n_v = 10$ parameters are set as independent variables, including power supply voltage, drive capacity of the drive buffer, buffer pad capacitance, Processor ODT value, motherboard breakout length, transmission line from via after series term, PCB trace length, DRAM pkg trace propagation delay, DRAM pad capacitance, etc. The rest $n_c = 35$ parameters are set as constants. We denote n parameters as $X = \{x_0, x_1, \ldots, x_n\}$. The model is to predict the accurate eye height y^{height} and width y^{width} metrics of the eye-diagram according to the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2019.2948527, IEEE Transactions on Circuits and Systems II: Express Briefs

TABLE I
PARAMETERS OF THE DNN NETWORK FOR HNN

Number of neurons	$L = \{45, 512, 128, 64, 32, 1\}$
Activation function	sigmoid function
Learning rate	0.005
Batch size	32
Maximum iterations	5000

TABLE II
PARAMETERS OF THE CNN NETWORK FOR HNN

Number m of segments for data	10
Input dimension	$(10 \times 45 \times 1 \times 1)$
Convolution kernel 1	$(3 \times 3 \times 1 \times 32)$
Convolution kernel 2	$(3 \times 3 \times 32 \times 64)$
Convolution kernel 3	$(1 \times 2 \times 64 \times 128)$
Size of fully connected layer	$L = \{1024, 512, 1\}$
Activation function	relu function
Learning rate	0.003
Batch size	32
Maximum iterations	5000

circuit parameters X of the system. Totally 2700 data are prepared as input. We generate 500 labeled data set L by using ICAT software. The rest 2200 unlabelled data formulate data set U. The eye-diagram metrics y_i^{height} ranges from $20.42mV$ to $225.95mV$, y_i^{width} ranges from $24.38ps$ to $138.91ps$.

B. Model Parameters Setting

The parameter settings of the DNN are shown in Table I. The parameter settings of the CNN are shown in Table II. The sampling rate γ of the HNN is set to 50%. The total number of iterations of DNN and CNN is 5000, which represents the number of iterations of HNN.

We use the root-mean-square error(RMSE) and the maximum relative error to measure the prediction performance. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(\hat{y}_i - y_i)^2}. \tag{13}$$

The relative error rate is defined as:

$$RE = |\hat{y}_i - y_i|/y_i, \tag{14}$$

The Maximum RE is defined as the maximum of the RE of all data.

C. Performance Compared with Traditional Method

As mentioned before, we extensively studied the performance and the characteristics of the Double-Data-Rate Synchronous Dynamic Random Access Memory (SDRAM) model by using the DNN method proposed in [12] and the COREG method proposed in [17]. We compare the performance of these two methods with our proposed HNN method. Each method has applied 500 labeled data. In DNN, 400 labeled data make up the training data set, while the remaining 100 make up the test data set. In HNN, 500 data sets with labels are divided into 3 parts, 200 data are used for input data set L_d for DNN, 200 data L_c for CNN, and 100 for public test set L_{te}. The unlabeled data set U is added on U_d and U_c to get the data set for HNN as:

$$D^H = \{L_d \cup U_d, L_c \cup U_c\}. \tag{15}$$

TABLE III
ACCURACY OF PREDICTED EYE HEIGHTS AND WIDTHS FROM THE HNN AND HNN

regression		height	width
DNN [12]	RMSE	11.86mV	6.14ps
	Maximum RE(%)	16.24	8.31
COREG [17]	RMSE	10.13mV	5.27ps
	Maximum RE(%)	9.68	7.16
HNN	RMSE	8.03mV	4.08ps
	Maximum RE(%)	4.61	5.06

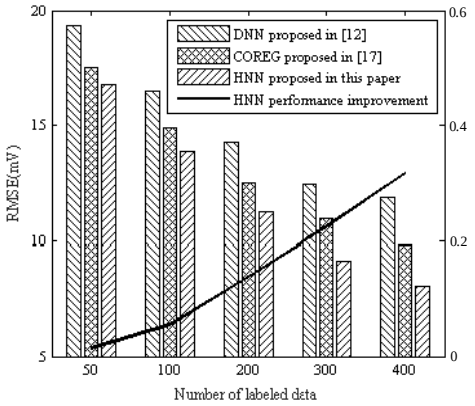

Fig. 5. Eye height prediction performance of HNN and DNN models in different labeled data set sizes

The accuracy of predicting eye height and width of HNN, DNN, and COREG is shown in Table III. According to Table III, compared with the DNN method proposed in [12], our result performance increases 32.29% in RMSE on height prediction, 33.55% in RMSE on width prediction, and 71.61% in Maximum RE on test data set. Meanwhile, compared with the COREG method, our method has improved the height prediction by 20.73% and the width prediction by 22.58%.

DNN uses only labeled samples for training, but the proposed HNN employs the semi-supervised learning method. As described in Section III, HNN uses the difference between DNN and CNN network to label unlabeled data reliably, increases the size of training samples, and achieves better performance under the same number of training samples.

Fig. 5 shows the RMSE of predicted eye height under a different number of labeled data. For instance, if our goal is to reach the accuracy of about 12mV RMSE, the proposed HNN only requires 200 training data while COREG requires 300 training data and DNN in [12] requires 400 training data to achieve the same accuracy. That is to say, HNN can save

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2019.2948527, IEEE Transactions on Circuits and Systems II: Express Briefs

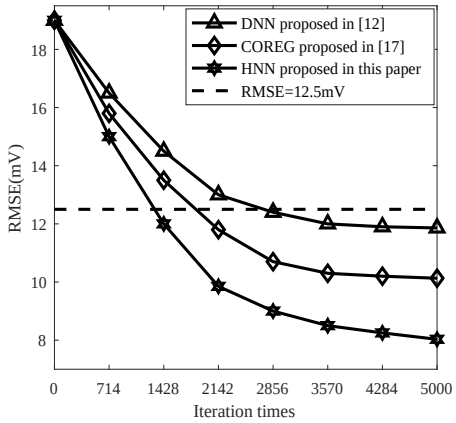

Fig. 6. The convergence of the training process of HNN and DNN on predicting eye height

TABLE IV
TIME COMPLEXITY OF VARIOUS METHODS

	ICAT	DNN [12]	HNN
Computing platform	Intel(R) core(TM) i7-8700 CPU GPU isn't supported	NVIDIA GTX 1080Ti i7-8700 CPU	NVIDIA GTX 1080Ti i7-8700 CPU
RMSE of height	True label	11.86mV	8.03mV
Time consumption	1h	1.6ms 2.7ms(only CPU)	3.5ms 6.4ms(only CPU)

almost 33.33% and 50% of the labeled data when it achieves the same performance as COREG and DNN.

As shown in Fig. 6, with only a few labeled data, HNN converges faster than DNN and COREG methods, and HNN can converge to a lower RMSE value.

As shown in TABLE IV, we compare the time consumption by HNN method with that by traditional simulation software and DNN method in predicting eye height. Experimental results show our method significantly cuts the time cost compared with the traditional signal integral analysis simulator.

IV. CONCLUSION

In this work, we propose a hybrid neural network for learning the Signal Integrity Model of high-speed chip-to-chip system. Only a few data with labels and reasonable amounts of unlabeled data are needed for training, compared with existing study that based on DNN solely. Once the learning is completed, the learned models can be used to predict the eye height and the eye width in future designs. The proposed method reduce 50% labeled data for training with 32.29% improvement on eye-diagram height prediction accuracy and 33.55% on eye-diagram width prediction. the proposed Semi-supervised learning-based approach requires no substantial domain knowledge nor massive amounts of labeled data thus saves complex and expensive circuit simulations.

REFERENCES

[1] R. Oikawa, T. Ochiai, T. Kida, K. Sakata, S. Kariyazaki, Y. Kayashima, Y. Ono, R. Mori, and T. Nomura. A high-speed, long-reach signal design challenge for 2.5-d lsi based on a low-cost silicon interposer and a large-scale sso analysis. In *2015 IEEE 65th Electronic Components and Technology Conference (ECTC)*, pages 293–300, May 2015.

[2] J. Fan, X. Ye, J. Kim, B. Archambeault, and A. Orlandi. Signal integrity design for high-speed digital circuits: Progress and directions. *IEEE Transactions on Electromagnetic Compatibility*, 52(2):392–400, May 2010.

[3] T. J. Brazil. Nonlinear, transient simulation of distributed rf circuits using discrete-time convolution. In *2007 IEEE International Symposium on Circuits and Systems*, pages 505–508, May 2007.

[4] S. Choi, H. Kim, K. Kim, J. Park, D. H. Jung, and J. Kim. Signal integrity analysis of silicon/glass/organic interposers for 2.5d/3d interconnects. In *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)*, pages 2139–2144, May 2017.

[5] D. Drogoudis, J. Van Hese, B. Boesman, and D. Pissoort. Combined circuit/full-wave simulations for electromagnetic immunity studies based on an extended s-parameter formulation. In *2014 IEEE 18th Workshop on Signal and Power Integrity (SPI)*, pages 1–4, May 2014.

[6] S. S. George, S. Sivanantham, S. Pawar, and R. Vikram. Signal integrity and power integrity challenges in embedded computing boards. In *2018 15th International Conference on ElectroMagnetic Interference Compatibility (INCEMIC)*, pages 1–4, Nov 2018.

[7] J. Chen, S. Chen, Q. Wang, B. Cao, G. Feng, and J. Hu. iraf: A deep reinforcement learning approach for collaborative mobile edge computing iot networks. *IEEE Internet of Things Journal*, 6(4):7011–7024, Aug 2019.

[8] Q. Xuan, B. Fang, Y. Liu, J. Wang, J. Zhang, Y. Zheng, and G. Bao. Automatic pearl classification machine based on a multistream convolutional neural network. *IEEE Transactions on Industrial Electronics*, 65(8):6538–6547, Aug 2018.

[9] Xiaoyong Gao, Chao Shang, Yongheng Jiang, Dexian Huang, and Tao Chen. Refinery scheduling with varying crude: A deep belief network classification and multimodel approach. *AIChE Journal*, 60(7):2525–2532, 2014.

[10] Xiao-Dong Li, J. K. L. Ho, and T. W. S. Chow. Approximation of dynamical time-variant systems by continuous-time recurrent neural networks. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 52(10):656–660, Oct 2005.

[11] Yong Fang and T. W. S. Chow. Nonlinear dynamical systems control using a new rnn temporal learning strategy. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 52(11):719–723, Nov 2005.

[12] 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*, 60(6):1957–1964, Dec 2018.

[13] Wenjian Zheng, Yi Liu, Zengliang Gao, and Jianguo Yang. Just-in-time semi-supervised soft sensor for quality prediction in industrial rubber mixers. *Chemometrics and Intelligent Laboratory Systems*, 180:36–41, 2018.

[14] Qi Xuan, Haoquan Xiao, Chenbo Fu, and Yi Liu. Evolving convolutional neural network and its application in fine-grained visual categorization. *IEEE Access*, 6:31110–31116, 2018.

[15] Qiuqiang Kong, Yong Xu, Iwona Sobieraj, Wenwu Wang, and Mark D Plumbley. Sound event detection and time–frequency segmentation from weakly labelled data. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 27(4):777–787, 2019.

[16] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the eleventh annual conference on Computational learning theory*, pages 92–100. Citeseer, 1998.

[17] Zhi-Hua Zhou and Ming Li. Semi-supervised learning by disagreement. *Knowledge and Information Systems*, 24(3):415–439, 2010.

[18] X. Liu, M. Masana, L. Herranz, J. Van de Weijer, A. M. Lpez, and A. D. Bagdanov. Rotate your networks: Better weight consolidation and less catastrophic forgetting. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 2262–2268, Aug 2018.

[19] M. Kashyap, K. Keshavan, and A. Varma. A novel use of deep learning to optimize solution space exploration for signal integrity analysis. In *2017 IEEE 26th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)*, pages 1–3, Oct 2017.

[20] D. Wang, M. Zhang, Z. Li, J. Li, C. Song, J. Li, and M. Wang. Convolutional neural network-based deep learning for intelligent osnr estimation on eye diagrams. In *2017 European Conference on Optical Communication (ECOC)*, pages 1–3, Sept 2017.