Enhancing Performance of Machine Learning-Based Modeling of Electromagnetic Structures

Zhou, Zhao; Wei, Zhaohui; Ren, Jian; Yin, Yingzeng; Pedersen, Gert Frølund; Shen, Ming

Published in:
2023 IEEE Conference on Antenna Measurements and Applications, CAMA

DOI (link to publication from Publisher):
10.1109/CAMA57522.2023.10352704

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

Enhancing Performance of Machine Learning-Based Modeling of Electromagnetic Structures

Zhao Zhou
Department of Electronic Systems
Aalborg University
Aalborg, Denmark
https://orcid.org/0000-0001-7895-8129

Zhaohui Wei
Department of Electronic Systems
Aalborg University
Aalborg, Denmark
https://orcid.org/0000-0003-4108-7686

Jian Ren
School of Electronic Engineering
Xidian University
Xi'an, China
https://orcid.org/0000-0001-9899-5963

Yingzeng Yin
School of Electronic Engineering
Xidian University
Xi'an, China
https://orcid.org/0000-0002-9103-9925

Gert Frølund Pedersen
Department of Electronic Systems
Aalborg University
Aalborg, Denmark
https://orcid.org/0000-0002-6570-7387

Ming Shen
Department of Electronic Systems
Aalborg University
Aalborg, Denmark
https://orcid.org/0000-0002-9388-3513

***Abstract*—The machine learning (ML)-based modeling of electromagnetic (EM) structures involves the development of a surrogate model that approximates the relationship between EM geometries and responses, such as S_{11}, gain, etc. The performance of the surrogate model is mainly affected by the simulation data for training. Normally, the training data is collected by uniformly sweeping the geometric parameters. Restricted by the computation power, only a limited parameter space can be sampled. The trained surrogate model behaves well within the sampling range but deteriorates as the parameter range extends. In this paper, we expand the predictable parameter range of an ML model with the same simulation expense by optimizing the data acquisition strategy. This approach leads to the proposed model demonstrating higher accuracy within an extended parameter space than conventional models, while the simulation consumption remains the same. We present an application example to validate its effectiveness. The proposed modified ML-based design method can potentially improve the performance of surrogate models in real-world applications.**

***Index Terms*—electromagnetic, machine learning, modeling, surrogate model**

I. Introduction

Electromagnetic (EM) structures (e.g., antennas, filters, metasurfaces, etc.) play an essential role in modern wireless communication systems. The modeling of EM structures means to determine the EM responses (e.g., Gain, S_{ij}, etc.) for a given setting of geometric parameters. It is ruled by Maxwell's theory, but can hardly be formulated quantitatively. The conventional modeling of EM structures relies on time-consuming and computation-expensive simulation software (e.g., Computer Simulation Technology (CST)). Machine Learning (ML) has been widely applied and validated for the fast modeling of EM structures by developing a surrogate model that approximates the projection between geometric parameters and EM responses [1]–[3].

ML-based modeling of an EM structure starts by sweeping its key geometric parameters over a specific parameter space and simulating all the combinations to corresponding EM responses. These pairs of geometric parameters and simulated EM responses are taken as training data to develop a surrogate model. The well-trained surrogate model will then be expected to approximate the projection of the EM structure: known a parameter setting, predicting its EM responses [2]; known EM responses, determining the parameter setting that leads to them [1]. The approximation accuracy and range heavily depend on the sweeping density and range. Provided a fixed computation budget, only a limited parameter space can be sampled at a certain sampling density to reach an expected accuracy within this parameter space. The approximation performance out of the sampling space deteriorates.

This paper enhances the performance of ML-based modeling of EM structures. We expand the predictable parameter range of an ML model with the same simulation expense by optimizing the data acquisition strategy. Instead of sweeping over a fixed parameter space uniformly, the training data are generated iteratively, and the parameter settings to collect are adjusted dynamically during each iteration based on the analysis of existing simulation data. The proposed approach investigates the correlation among data samples and optimizes the training data distribution to maximize the informativeness of the collected training data set while maintaining the simulation expense. The proposed method is implemented on a microwave filter to validate its effectiveness.

II. Method and Validation

The proposed method is implemented on a microwave filter proposed by Yang *et al.* [4] for validation. It operates from $0.1\,\mathrm{GHz}$ to $4.5\,\mathrm{GHz}$. Fig. 1 exhibits the structure of the microwave filter, which consists of three metal layers and two substrate layers ($Rogers RO4003C$ with relative permittivity of $\varepsilon_r = 3.38$ and loss tangent of $tan\delta = 0.0027$). The top and bottom metal layers are open-ended two-order microstrip feedlines are centrosymmetric to each other. The middle layer

This paper is in part supported by the China Scholarship Council.

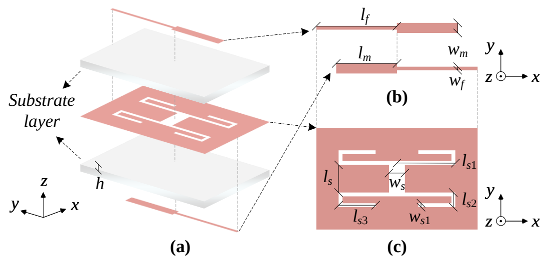

Fig. 1. Structure of the microwave filter [4]. (a) 3D view; (b) Top and bottom metal layers; (c) Middle metal layer.

TABLE I
GEOMETRIC PARAMETERS OF THE MICROWAVE FILTER [4]

Parameter	Value (mm)	Parameter	Value (mm)
l_{s1}	[11.5, 12.5]	w_f	1.5
l_{s2}	[2, 3]	w_s	1.5
l_{s3}	[9.5, 10.5]	w_{s1}	0.5
l_m	20.5	w_m	3.88
l_s	9.9	h	0.813
l_f	24		

TABLE II
TRAINING AND TESTING DATA OF CONVENTIONAL METHOD

Parameter	Training (343 samples)			Testing (988 samples)		
(mm)	min	max	step	min	max	step
l_{s1}	11.7	12.3	0.1	11.5	12.5	0.1
l_{s2}	2.2	2.8	0.1	2	3	0.1
l_{s3}	9.7	10.3	0.1	9.5	10.5	0.1

Note: Testing samples exclude 343 training samples.

TABLE III
ARCHITECTURE OF THE MACHINE LEARNING MODEL

No.	Layer	Neurons	Function
1	Input	3	Input: [l_{s1}, l_{s2}, l_{s3}]
2	Activation function	-	Tanh
3	Fully-connected	160	Hidden layer
4	Activation function	-	Tanh
5	Fully-connected	180	Hidden layer
6	Activation function	-	Tanh
7	Output	45	Output: [$\|S_{11}\|$]
8	Activation function	-	Linear
	Loss function		Mean Squared Error
	Optimizer		Adam
	Learning rate		0.001
	Batch size		10
	Epochs		180

is a rectangular patch etched with a symmetric four-folded-branched slot. The geometric parameters are marked in Fig. 1 and listed in Table I. Three geometric parameters (l_{s1}, l_{s2}, l_{s3}) are selected as adjustable variables, while the rest parameters are fixed as constant values.

A forward model is defined to model the microwave filter. It aims to predict the $|S_{11}|$ for a given setting of l_{s1}, l_{s2}, and l_{s3}. Each $|S_{11}|$ is discretely sampled and represented as a 45-element vector. Accordingly, the input size of the surrogate model is 1 by 45, and its output size is 1 by 3. The forward model is first developed from scratch by using the conventional method. Its training and testing data are listed in Table II. The size of testing data is larger than the training data because the testing data is an expanded set of the training data. The mean squared errors (MSE) on the training set and testing set are 6.9×10^{-4} and 1.8×10^{-3}, respectively. It shows that the model performance deteriorates significantly out of the sampling parameter space.

We then utilize the proposed method to develop the forward model from scratch:

1) Sweep l_{s1}, l_{s2}, and l_{s3} over [11.7, 12.3], [2.2, 2.8], and [9.7, 10.3] on a sparse and uniform sampling grid of $4 \times 4 \times 4$. Each parameter setting is normalized and formulated as a 3-element vector $\boldsymbol{X}$, and corresponding $|S_{11}|$ is normalized and formulated as a 45-element vector $\boldsymbol{Y}$.

2) Pick a reference input vector $\hat{\boldsymbol{X}}_j$ for each existing input vector $\boldsymbol{X}_j$, by minimizing the distance between $\hat{\boldsymbol{X}}_j$ and $\boldsymbol{X}_j$. Form a list of reference samples $\{(\hat{\boldsymbol{X}}_j, \hat{\boldsymbol{Y}}_j)\}_{N_t}$ with respect to existing samples $\{(\boldsymbol{X}_j, \boldsymbol{Y}_j)\}_{N_t}$.

3) Pick one input vector $\boldsymbol{X}_k$ by maximizing the distance between its output $\boldsymbol{Y}_k$ and its reference output $\hat{\boldsymbol{Y}}_k$. The selected sample ($\boldsymbol{X}_k$, $\boldsymbol{Y}_k$) and its reference sample ($\hat{\boldsymbol{X}}_k$, $\hat{\boldsymbol{Y}}_k$) point at an input space where the sample of high quality exists. The underlying reason is that selected samples have a large distance between their output vectors and a small distance between their input vectors, implying that they confine a parameter space where the output is sensitive to the input. Hence, it is likely that this space contributes a lot to the prediction error of the ML model. On the other hand, adding a new sample in this space can improve the model performance significantly. In this sense, a potential high-quality sample of high informativeness referred to as ($\boldsymbol{X}_*$, $\boldsymbol{Y}_*$) can be generated from the selected sample ($\boldsymbol{X}_k$, $\boldsymbol{Y}_k$) and its reference sample ($\hat{\boldsymbol{X}}_k$, $\hat{\boldsymbol{Y}}_k$).

4) Examine if the absolute difference between $\boldsymbol{X}_k$ and $\hat{\boldsymbol{X}}_k$ exceeds $2 \times \boldsymbol{X}_{step}$. If yes, it guarantees an input space large enough to generate a new input vector; otherwise, go to step 3) after excluding these two samples.

5) Train the forward model using the existing training data. Check its MSE on the training dataset (343 samples). If its MSE is higher than 6.9×10^{-4}, go to step 6); otherwise, go to step 7).

6) Generate a new input vector $\boldsymbol{X}_*$ from the selected input vector $\boldsymbol{X}_k$ and its reference vector $\hat{\boldsymbol{X}}_k$ through formula (1):

$$\boldsymbol{X}_* = 0.5 \cdot \boldsymbol{X}_k + 0.5 \cdot \hat{\boldsymbol{X}}_k. \tag{1}$$

The input vector of the new sample is obtained through swarm operation, as given in formula (1), on the selected input vector $\boldsymbol{X}_k$ and its reference input vector $\hat{\boldsymbol{X}}_k$. Here, the input vector of the new sample is referred to as $\boldsymbol{X}_*$ and it is set as the weighted sum of the $\boldsymbol{X}_k$ and $\hat{\boldsymbol{X}}_k$. The weights for $\boldsymbol{X}_k$ and $\hat{\boldsymbol{X}}_k$ is 0.5. Go to step 8).

7) Check if $\boldsymbol{X}_k$ has one of the minimum or maximum values (11.7, 2.2, 9.7, or 12.3, 2.8, 10.3). If not, go to step 3) after excluding these two samples. Otherwise, generate a new input vector from the two samples. The generation follows equation 2,

$$\boldsymbol{X}_* = 2 \cdot \boldsymbol{X}_k - \hat{\boldsymbol{X}}_k. \tag{2}$$

TABLE IV
COMPARISON OF PERFORMANCE BETWEEN PROPOSED METHOD AND CONVENTIONAL METHOD

MSE	Sampled space	Expanded space
Conventional method	6.9×10^{-4}	1.8×10^{-3}
Proposed method	6.9×10^{-4}	1.1×10^{-3} ($\downarrow 0.7 \times 10^{-3}$)

Go to step 8).

8) Obtain the output label vector $\boldsymbol{Y}_*$ for this new input vector $\boldsymbol{X}_*$ through full-wave simulation via CST.

9) Form a new sample ($\boldsymbol{X}_*$, $\boldsymbol{Y}_*$), add this new sample to the existing data set, and repeat sub-steps (a-h) until sufficient data have been acquired.

The optimized dataset with the same size of 343 data samples is used to train the forward model. The well-trained model is evaluated on the same training and testing dataset in Table II again. The training and testing MSEs are 6.9×10^{-4} and 1.1×10^{-3}. Compared to conventional methods, as shown in Table IV, the proposed method shows superior modeling performance in the expanded parameter space while maintaining accuracy within the sampled parameter space.

III. CONCLUSION

This paper proposes a modified machine learning (ML)-based modeling method for electromagnetic (EM) structures. The proposed method enhances the modeling performance by optimizing the data acquisition strategy. A microwave filter is utilized to validate its effectiveness. The comparative results demonstrate that the proposed method enhances the modeling performance in an expanded parameter space significantly compared to conventional methods.

REFERENCES

[1] M. Sedaghat, R. Trinchero, Z. H. Firouzeh, and F. G. Canavero, “Compressed machine learning-based inverse model for design optimization of microwave components,” *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 7, pp. 3415–3427, Apr. 2022, doi: 10.1109/TMTT.2022.3166151.

[2] S. Koziel and M. Abdullah, “Machine-learning-powered EM-based framework for efficient and reliable design of low scattering metasurfaces,” *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 4, pp. 2028–2041, Mar. 2021, doi: 10.1109/TMTT.2021.3061128.

[3] Z. Wei, Z. Zhou, P. Wang, J. Ren, Y. Yin, G. F. Pedersen, and M. Shen, “Equivalent circuit theory-assisted deep learning for accelerated generative design of metasurfaces,” *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5120–5129, Feb. 2022, doi: 10.1109/TAP.2022.3152592.

[4] L. Yang, L. Zhu, R. Zhang, J. Wang, W. Choi, K. Tam, and R. Gómez-García, “Novel multilayered ultra-broadband bandpass filters on high-impedance slotline resonators,” *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 1, pp. 129–139, Jan. 2019, doi: 10.1109/TMTT.2018.2873330.