

Data Augmentation for Optimization of Electromagnetic Devices Using Machine Learning

Shuli Yin, Hajime Igarashi

Graduate School of Information of Technology, Hokkaido University, Japan, shuli.yin@polimi.it, igarashi@ssi.ist.hokudai.ac.jp

This paper proposes a scheme of applying data augmentation to improve the accuracy of machine learning for optimal design of electromagnetic devices, where an inductor design with neural network (NN) is considered as a concrete example. The data set is augmented by introducing perturbative noise to geometric parameters, while keeping the inductance, core loss, and magnetic saturation capacity to their original values. Through K-fold cross-validation and Pearson correlation coefficient, it has been demonstrated that the NN model with data augmentation achieves higher accuracy than the one without data augmentation.

Index Terms— Data augmentation, perturbations, regression, three-dimensional inductor design.

I. INTRODUCTION

IN DESIGN of electrical components or devices using optimization techniques such as evolutionary and heuristic algorithms, a large number of searches are required. In this process, the most time-consuming step is the simulation evaluation. Especially in the design of inductors, numerical computational methods such as finite element method (FEM) are usually implemented to obtain the performance metrics included in the objective function, which significantly reduces the search efficiency.

Neural network (NN) has been proved to be very successful in the field of regression tasks [1], where the goal is to predict continuous output values based on input data. By leveraging the surrogate models for inductors using NN, FEM can be substituted, thus greatly compressing optimization time.

Although NN is powerful in improving efficiency for regression problems, prediction accuracy is also a key aspect to consider. The prediction accuracy of a trained NN model is highly dependent on the data, and insufficient data may lead to overfitting or underfitting. To mitigate these problems, data augmentation, which artificially enriches the data, would be effective. Despite extensive research and studies on data augmentation, most of the focus has been on image classification and recognition [2]. The data augmentation for regression has only been studied for time series forecasting [1].

In this study, the data augmentation is applied to the regression for performance prediction of three-dimensional (3D) inductor models using NN. The originality lies in the fact that this is the first time for applying data augmentation to the optimal designs of electromagnetic devices.

II. METHODOLOGY

Let us consider a 3D inductor model including serially connecting the upper and lower spiral coils with rectangular section surrounded by magnetic core [3]. The structure of coils and the material properties of the core are shown in Fig. 1. Key geometric variables which influence the electromagnetic characteristics of the inductor, such as the radius r , coil width w , thickness t , etc., are also marked in Fig. 1. These variables are represented by $\mathbf{x}_{ori} = [h, t, p, w, r, l]$, which affects the characteristics of the inductor such as the inductance value (denoted by L_{1A}) when the input current I_{in} equals 1 A, the iron

loss P , and the saturation capability of the inductor, which is here characterized by F defined as the ratio of the inductance at $I_{in} = 18$ A (denoted by L_{18A}) to L_{1A} . These characteristics are represented by $\mathbf{y}_{ori} = [L_{1A}, P, F]$. We aim to train a NN model with $\mathbf{x}_{ori}$ as the input and $\mathbf{y}_{ori}$ as the output.

Before training, the data need to be prepared. The range of $\mathbf{x}_{ori}$ is summarized in TABLE I. Within this range, 641 distinct sets of $\mathbf{x}_{ori}$ are randomly selected, and the corresponding $\mathbf{y}_{ori}$ is computed by FEM. To enhance the model robustness, the normalization on data is then performed, $\mathbf{x}_{ori}$ and $\mathbf{y}_{ori}$ are transformed to be $\mathbf{x}$ and $\mathbf{y}$, which yield

$$\mathbf{x} = (\mathbf{x}_{ori} - \min(\mathbf{x}_{ori})) / (\max(\mathbf{x}_{ori}) - \min(\mathbf{x}_{ori})), \quad (1a)$$

$$\mathbf{y} = (\mathbf{y}_{ori} - \min(\mathbf{y}_{ori})) / (\max(\mathbf{y}_{ori}) - \min(\mathbf{y}_{ori})). \quad (1b)$$

Then the NN consisting of two hidden layers is constructed as illustrated in Fig. 2. The first hidden layer contains 64 neurons ($L_1 = 64$), and the second hidden layer contains 32 neurons ($L_2 = 32$). Labels $w_{i,j}$, $w'_{i,j}$, and $w''_{i,j}$ ($i=1, 2, \dots, 6$, and $j=1, 2, \dots, L_1$ or L_2) represent the weights between adjacent layers; $b_{i,j}$, $b'_{i,j}$, and $b''_{i,j}$ corresponds to bias. Through adjusting weights and bias, the global error between the actual output and expected output is minimized. Then the predicted values $\hat{\mathbf{y}}$ are obtained using the trained model. After applying the inverse normalization, the predicted values in the original scale are finally recovered.

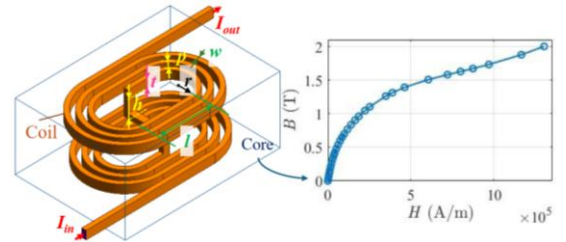


Fig. 1. Illustration of the 3D inductor, and the nonlinear performance of the core.

TABLE I
RANGE OF GEOMETRIC VARIABLES

Item	Min. (mm)	Max. (mm)
h	0.04	0.50
t	0.04	0.40
p	0.12	0.20
w	0.01	0.10
r	0.20	0.60
l	0.10	0.80

The above outlines the classical training process for a NN model. To implement the data augmentation, after the normalization, several sets of data are randomly selected for adding noise δ , which is assumed to lie within the range $[-0.1, 0.0) \cap (0.0, 0.1]$. When randomly selecting data sets from the training data, new data sets can be generated by adding perturbative noise to several feature parameters of $\mathbf{x}$, i.e.,

$$\mathbf{x}_i = \mathbf{x}_i + \delta, \quad \text{if } i = 1, 2, \dots, 6 \quad (2)$$

while keeping corresponding $\mathbf{y}$ unchanged assuming that δ is sufficiently small. Through this approach, the data set is expanded to 5,000 in this work. Additionally, in this study, it is limited that no more than 4 feature parameters can contain noise in a single data set. Then the subsequent process follows the same training procedure as that of the classical one mentioned above.

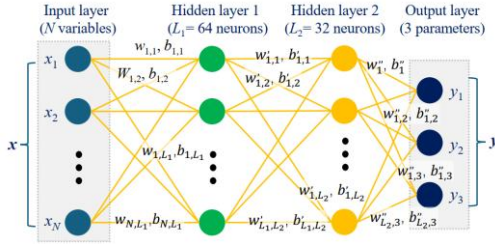


Fig. 2. Neural network architecture.

III. RESULTS AND COMPARISONS

To evaluate the performance of the NN model, five-fold cross-validation is employed. To be more precise, datasets are randomly split into five groups of the same size. During each validation, one fold is used as the validation data, while the other four groups are combined and used as the training data. Then performance of the NN model can be evaluated by repeating this process until all groups have been validated. It is worthy mentioning that in each validation, only the training data are used for augmentation.

Then the correlations between the actual values and the predicted ones for each validation can be finally obtained, as shown in Fig. 3 (a). For comparison, the correlations using the classical NN model are also given, as shown in Fig. 3 (b). It can be observed that the results align better with the true values using data augmentation to the classical one.

To quantify this, the L_2 -norm relative errors of the predicted values to the true values are also evaluated, as listed in TABLE II. The results demonstrate that, in each validation step (designated as Trial 1 to Trial 5) including the five-fold average value, errors obtained from the augmented-data NN model, are smaller than those obtained using the classical one.

Furthermore, TABLE II also lists the L_2 -norm errors when the low noise (LN) is applied to augment data into 5,000 sets, i.e., $\delta \in [-0.005, 0.0) \cap (0.0, 0.005]$. It is evident that when the noise is negligible, increasing the training data volume does not improve the accuracy of the NN model. Therefore, non-negligible perturbations are an essential factor for accuracy.

Except cross-validation, the performance of the model is also assessed through Pearson correlation r_k , which reads

$$r_k = \sum_{\text{Trial}=1}^5 [\text{cov}(\mathbf{y}_k, \hat{\mathbf{y}}_k) / (\sigma_{\mathbf{y}_k} \cdot \sigma_{\hat{\mathbf{y}}_k})] / 5 \quad (3)$$

where k represents a feature included in $\mathbf{y}$; $\sigma_{\mathbf{y}_k}$ and $\sigma_{\hat{\mathbf{y}}_k}$ are the standard deviation of the the actual and the predicted results for the k -th feature. A higher Pearson correlation coefficient suggests the predictions of NN closer aligning with the true values. Upon calculation, we obtain $r_{L1A} = 0.9845$, $r_P = 0.9836$, and $r_F = 0.9774$ for the classical model, while the models with augmented data yield 0.9955, 0.9950, and 0.9911, respectively, clearly indicating that the latter is more consistent with the true results.

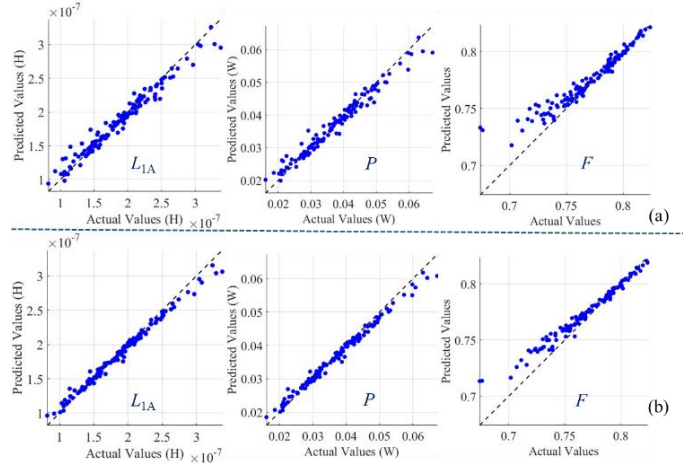


Fig. 3. Actual vs prediction plots for one trial of the validation. (a) classical NN model, (b) augmented-data NN model.

TABLE II
 L_2 -NORM RELATIVE ERRORS

Item	L_{1A} (%)			P (%)			F (%)		
	Cls.	Aug.	LN	Cls.	Aug.	LN	Cls.	Aug.	LN
Trial 1	8.84	8.16	8.58	9.29	8.48	9.48	1.36	1.09	1.37
Trial 2	8.68	8.22	8.75	9.84	9.27	10.02	1.69	1.53	1.65
Trial 3	4.62	3.12	4.82	5.32	3.25	5.33	1.20	1.09	1.23
Trial 4	7.92	7.21	7.82	9.37	6.89	7.52	1.15	0.89	1.11
Trial 5	5.91	3.84	5.22	5.45	3.95	6.22	1.33	1.21	1.39
Ave.	7.19	6.11	7.04	7.45	6.37	7.72	1.35	1.16	1.35

IV. CONCLUSION

This paper proposes the augmented-data NN model. Compared with the classical one, it is more accurate for predicting the performance of the 3D inductor. More details and application to other devices will be provided in the full paper.

REFERENCES

- [1] S. Demir, K. Mincev, K. Kok, et al, "Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting," *Applied Energy*, 304, 2021.
- [2] C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, *Journal of big data*, 6(1), 1-48, 2019.
- [3] Y. Sato, K. Kawano, H. Igarashi, Direct inverse modeling for electromagnetic components using Gaussian kernel regression, *IEEE Trans. Magn.*, vol. 58, no.5, 8000708, 2022.