Computer-Aided Measurement method of Hysteresis Loop based on Convolution Neural Network

Ruohan Gong, Abdelkader Benabou, Zuqi Tang
Univ. Lille, Arts et Metiers Paris Tech, Centrale Lille, HEI, EA 2697
L2EP – Laboratoire d'Electrotechnique et d'Electronique de Puissance, F-59000 Lille, France

ABSTRACT – In this paper, a surrogate model base on deep learning (DL) is proposed to predict the hysteresis loops of ferromagnetic materials. The databases of hysteresis loops were measured on the MPG200D Brockhaus equipment with an Epstein frame. In order to reduce the experimental cost and accelerate measurements, a surrogate model based on the convolutional neural network (CNN) is proposed. First, presenting the measurement results in the form of 256$\times$ 256 $\times$ 1 images and extract the 6 most characteristic parameters, namely peak magnetic flux density, frequency, maximum magnetic field strength, remanence, coercivity, and the area of the hysteresis loop. All these physical parameters are taken as the label in the supervised DL process. These label information are normalized to form a Gaussian distribution image of 256$\times$ 256 $\times$ 3 as the input, and the corresponding B-H curve is the output. Using image-to-image CNN U-net, once the network is effectively trained, the hysteresis loops under other excitation parameters can be predicted without further measurement. Our numerical examples show that the prediction results agree well with the measurement results. The sensitivity of CNN for hysteresis loops prediction with respect to the hyperparameters are investagted. A set of empirical hyperparameter configurations is put forward to guarantee an efficient convergence. This research shows that the proposed approach can be an efficient tool to predict the hysteresis loop of ferromagnetic materials under different circumstances, which can potentially contribute to nonlinear hysteresis FEM computation.

Keywords – Convolutional neural network, deep learning, ferromagnetic materials, hysteresis loop.

1. INTRODUCTION

Ferromagnetic materials play an irreplaceable role in industrial technology. It has been widely used in navigation, machinery, medical and other fields. Reasonable and valid application of ferromagnetic materials requires knowing the detailed physical parameters, especially the ferromagnetic properties. In engineering practice, we usually obtain the required parameters by measuring the magnetization curve, hysteresis loop. Due to the wide application field and the particularity of ferromagnetic materials, we are often forced to make multiple measurements of the same group of material at different frequencies, magnetic flux densities, exciting currents and so on [1]. However, lots of experimental measurements can be time-consuming and also very expensive due to the requirements of specific test samples as well as complex measurement systems. This has brought some obstacles to the application of ferromagnetic materials.

On the other hand, although there is not yet a solid theoretical framework, the emergence and popularity of DL offer us a brand new approach. Over the last decade, DL has become an unprecedented tool that can essentially improve our capability to carry out scientific research due to advances in theory (solvers and optimizers) and infrastructure (larger memory and faster graphic processing units) [2]. CNN has gained tremendous popularity and has been widely used because of its ability to automatically capture high-level representative features of image, especially when it comes to computer vision [3, 4]. Deep learning, in particular, CNN has conquered many fields and achieved remarkable performance, especially in computer vision. Thanks to its sparse connectivity and shared weights, CNN can extract high-dimensional features from image data, especially effective when the data size is large. This makes DL a possible method to solve tasks involved with a large number of variables and complicated relationships [2]. We have done some previous work in this field, including applied CNN U-net in the optimal mesh refinement for the NDT problem [5] and the magento-thermal coupled analysis for the transformer [6] and got some satisfactory performances.

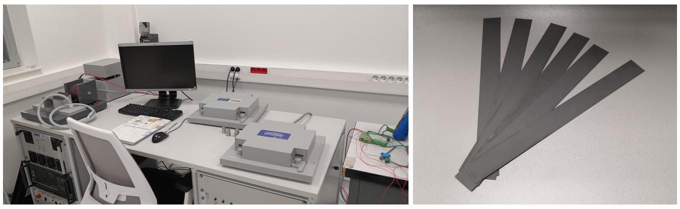
In this paper, a surrogate model based on CNN U-net is put forward to accelerate the hysteresis loops measurement of ferromagnetic materials. Based on a certain number of experimental measurements, these results are translated into pictures and labeled with the corresponding parameters. From the measurement data, two datasets are considered. The first one is used for the training samples to train the neural network and update the network weights iteratively. The rest of the other results are used for the cross-validation procedure to validate the effectiveness of the trained CNN model. The prediction results agree well with the measurement results, it is shown that the proposed method can be used as an efficiency tool to predict the hysteresis loop of ferromagnetic materials under different circumstances. Besides, based on measurement results and numerical experiments, we put forward a set of "empirical optimal" hyperparameters baseline configuration that can ensure the effective converge of the CNN. On this basis, we further conduct a sensitivity analysis of CNN U-net to explore the effects of hyperparameters on the model.

2. MEASUREMENT AND METHODOLOGY

Any forms of surrogate models require a certain number of the dataset so it can be effectively trained. To come up with reliable models, one has to have reliable data: the dataset should be able to represent all the situations that the surrogate model is intended for. In this scenario, we take the hysteresis loops obtained by experimental measurements as ground truth, part of these samples were used to train the model, the rest of the other samples are used to verify the accuracy of the model through cross-validation. After effective training of the model, the hysteresis loops under other circumstances within the parameter range of the database can be predicted by interpolation.

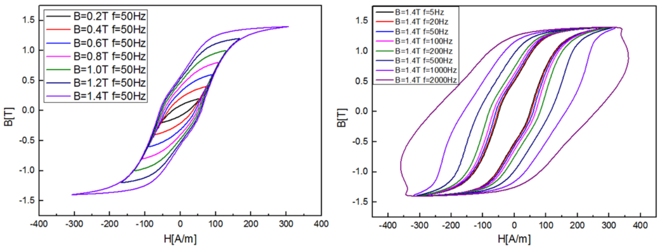
2.1. *Measurement results*

The measurements were performed on the MPG200D Brockhaus equipment with an Epstein frame as illustrated in Figure 1(a) and Figure 1(b). Hysteresis loops were measured at different frequencies and different peaks of magnetic flux density: the frequency ranges from 5 Hz to 2000 Hz, and the peak of magnetic flux density ranges from 0.1T to 1.8T with a step of 0.1T. All subsequent network training and learning process are limited to the range of the above parameters. The measurement results are illustrated by the hysteresis loops in Figure 2(a) and Figure 2(b). Finally, the measurement represents 532 sets of data, i.e. hysteresis loops, that constitute the database for the present study.



(a) (b)

Figure 1. Measurement device: (a) Brockhaus MPG200D equipment and (b) Epstein samples



(a) (b)

Figure 2. Measurement results under different excitation parameters: (a) varying magnetic flux density and (b) varying frequency

2.2. *Deep learning methodology*

Unlike traditional computer vision problems that have little rules that can be formulated and summarized such as object classification, satellite identification, etc. Problems in the field of electrical engineering are usually physics-based, which means that their behaviors are subject to strict physical rules. This means that there are bound to be some key parameters closely related to the final outcome. This facilitates the labeling of samples and the application of supervised learning. All these can help us to complete the network training more efficiently.

The schematic diagram of the proposed DL surrogate model for hysteresis loop prediction of ferromagnetic materials is described in Figure 3. The general process can be summarized as follows: To begin with, we translate experimental results into $256\times 256 \times 1$ images and label them with the corresponding characteristic parameters, i.e. peak magnetic flux density, frequency, maximum magnetic field strength, remanence, coercivity, and the area of the hysteresis loop. Then, the labeled database is divided into two parts. Using one part of labeled images as training samples, the weights of the network can be updated iteratively. The rest part is used for the cross-validation procedure to validate the effectiveness of the surrogate model. Once the CNN network is effectively trained, it can be used to predict the field distributions in other cases with only label information, thus accelerating the measurement process and saving experimental time. Taking these hysteresis loops images with denoted physical labels as the database, we introduce the CNN U-net described in [6] for image-to-image training in this study. The input is a 3-layer Gaussian distribution diagram composed of the information of 6 labels ane the output is the corresponding B-H curve. The Gaussian distribution will be converted to a $256\times 256 \times 3$ image and used as input to the network. Meanwhile, the output target is the $256\times 256 \times 1$ hysteresis loop corresponding to the physical labels. In addition, it is necessary to normalize the information of each label before generating the Gaussian distribution.

We choose U-net, which has demonstrated great potential in previous work, to explore the hidden relationship between the hysteresis loop and physical labels. This network has been successfully applied in [6] for the evaluation of magneto-thermal coupled analysis for a power transformer and in [5] for the optimal mesh refinement for the non-destructive testing (NDT) problem. U-net is a so-called “full convolutional network” that entirely composed of convolution and up-convolution. More detailed information about the U-net model can be found in [6]. The adaptive moment estimation (Adam) optimization algorithm is adopted to update the weights of the neural network [7, 8]. The mean absolute error (MAE) is used to evaluate the prediction results.

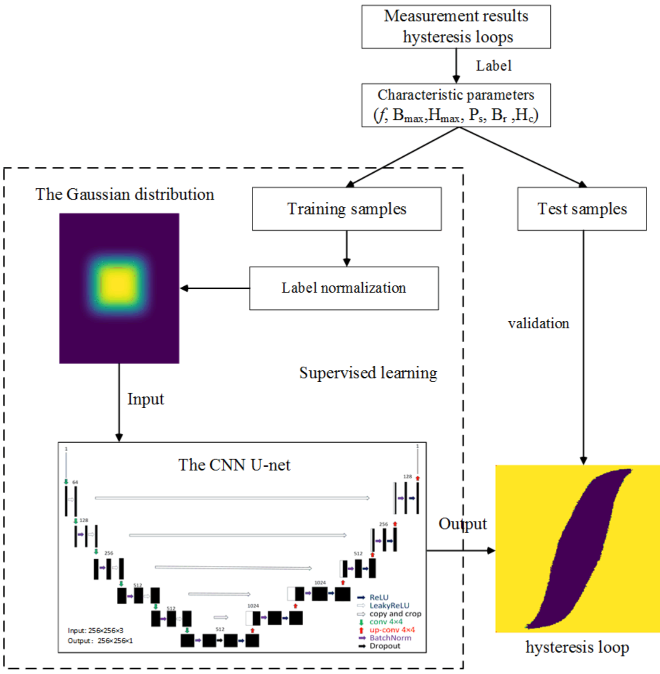


Figure 3. Schematic diagram of the surrogate model for hysteresis loop prediction

3. RESULTS AND SENSITIVITY ANALYSIS

Although DL has achieved some performances in several fields so far, it is still a well-known fact that the performance of DL deeply depends on the richness and diversity of the dataset. On the contrary, too many data sets can also put a heavy burden on experimental testing. In addition, the specific learning effect of the neural network is also affected by the network architecture and hyperparameters that to be tuned for each problem respectively. Considering that the purpose of the proposed DL surrogate model is to accelerate the measurement and reduce the cost, we need to investigate the scope and limitations of the model so that we can determine whether the approach can meet our needs. In the process of DL, besides constructing network architecture, the most important work is to specify these accompanying hyperparameters [9]. In this work, our goal is to identify a set of "possible best" empirical hyperparameters for hysteresis loop prediction and provide a reasonable range for each hyperparameter. This can provide reference for future research. So far, very little empirical data is available to guide such decisions.

3.1. *Baseline configuration*

There are many hyperparameters such as label normalizations, training sample size, learning rate, and batch size need to be determined so we can efficiently train the neural network. In practice, optimizing all these hyperparameters synchronously is simply not possible, especially when one hyperparameter will interact with many other aspects of the optimization process, not to mention that the interactions may be nonlinear. We first consider the performance of a baseline CNN U-net configuration.

To be specific, we start with the architectural decisions and hyperparameters adopted in previous work [6] and described in Table 1. Next, we will conduct the sensitivity analysis on each hyperparameter, and the other hyperparameters remain unchanged during the analysis.

Table 1. Hyperparameters of the baseline configuration

Hyperparameters	Value
Label normalization	Yes
Number of training samples	213
Learning rate	1e-4
Batch size	4

3.2. Label normalization

Different from the geometric variables that we were dealing with in the previous work, the physical labels in the hysteresis loop all have different units and their values fluctuate over a wide range. For instance, the frequency ranges from 5 Hz to 2000 Hz, and the peak of magnetic flux density ranges from 0.1T to 1.8T, by contrast, the maximum magnetic field strength ranges from 30 A/m to 8800 A/m, what's more, the area of hysteresis loop floats over a wide range from 0.001 W/kg to 330 W/kg. In this case, there must be some information loss if we directly combining original physical information into the Gaussian distribution. To address this problem, we preprocess the physical label in advance to normalize the values of the different labels to between 0 and 1.

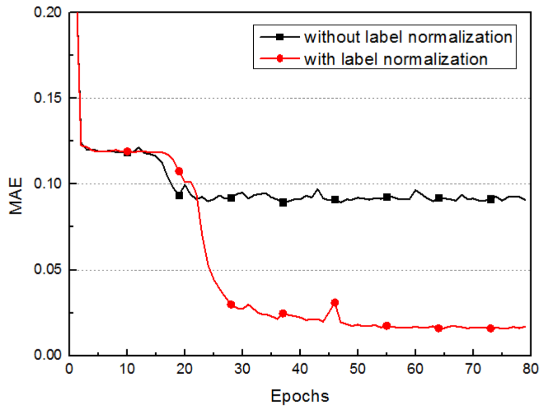


Figure 4. Training curve with or without label normalization

The effect of label normalization is illustrated in Figure 4. Obviously, for the same network architecture and hyperparameters, the network performance is greatly improved after the introduction of normalization. The MAE greatly dropped from 0.1 to about 0.02.

3.3. Number of training samples

According to some of our previous research in [5, 6], the most dominant hyperparameters are generally training sample size, learning rate and batch size. Hence, we will focus on these three factors with the greatest impact and conduct the sensitivity analysis in this work. The original objective of this paper is to accelerate the hysteresis loops measurement and reduce experimental cost. The proposed approach can reduce the number of measurements by predicting B-H curve without measurement once the model has been successfully trained by part of the ground truth data. The necessary condition to get a successful trained network is that enough training samples are available. The sample size is critical for learning effectiveness. Unfortunately, some DL models only work well when the data sets are very large, often over tens of thousands of samples [10]. In this case, the preparation of a large number of hysteresis loop test data itself is very time-consuming and laborious, which makes our acceleration auxiliary model meaningless.

From this aspect, given our previous work experience, U-net is a very powerful model when it comes to this situation. One of its greatest advantages is that it can be effectively trained when only a few training samples are available. Due to its excellent capacity for data augmentation and segmentation, U-net can extract the implicated patterns and associations between the inputs and outputs. In this case, there are 532 samples of B-H curve images, with 6 physical labels for which sample. 532 samples were divided into 10 groups from #0 to #9, with 53 or 52 samples in each group. Among which, some groups were used as training samples, and the rest groups are taken as test samples. The performance of the proposed surrogate model, in terms of hysteresis loops prediction, is shown in Figure 5 for different sizes of training samples.

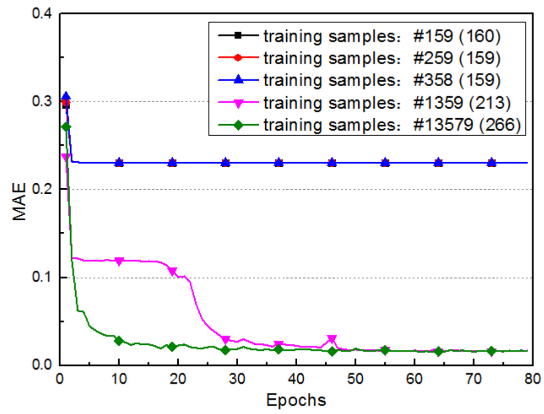


Figure 5. Training curves with different sizes of training samples

As we can seen in Figure 5. The proposed model can be effectively trained with only 40%–50% of the available samples. This makes the model an effective tool to reduce the experimental cost as well as the required memory for experimental data storage. The trained model can output the rest 256 prediction results within 1 minute. The proposed surrogate model has been proven to be effective in accelerating hysteresis loops measurement of ferromagnetic materials.

3.4. Learning rate

In addition to the proper network architecture and sufficient database, the CNN U-net still requires practitioners to set accompanying hyperparameters, including the learning rate and batch size. The learning rate is the amount of change to the model during each step of the iterative process, it controls the speed at which the model learns. Generally, a large learning rate allows the model to learn faster, with the cost of a more oscillating convergence process. A smaller learning rate may allow the model to learn a more optimal set of network weights but may take much longer to converge. A perfectly configured learning rate is essential to an effective training process.

As shown in Figure. 6, the optimal learning rate for B-H curve prediction is about 2e-5–5e-4. Inappropriate learning rates will lead to invalid training.

3.5. Batch size

The batch size controls the number of training examples utilized in one update to the network weights. An epoch is the entire training data exposed to the network, batch-by-batch. In general, the larger the batch size, the more accurate the decent gradient, but with that comes a significant increase in the time-consuming and storage requirement of one update. By contrast, the smaller the batch size, the more instability in gradient calculation, the network is harder to converge.

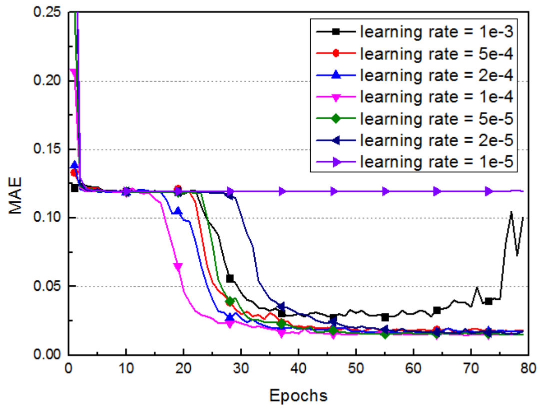


Figure 6. Training curves with different learning rates

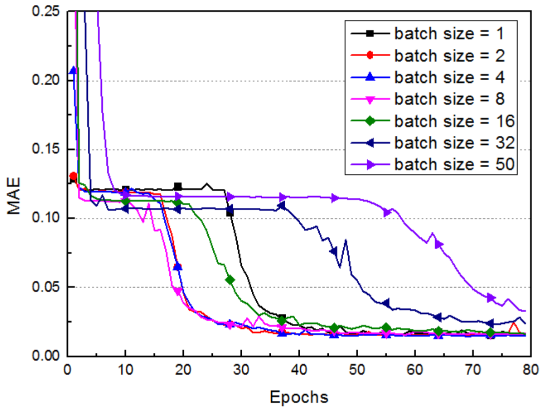


Figure 7. Training curves with different batch sizes

Like the learning rate, too large or too small the batch size will lead to invalid training. For the prediction of the hysteresis loops, the optimal batch size is about 1–16. The out-of-range batch size will affect the efficiency of the converging process. With the proper network parameters, the learning process can be fast and effective, even with only a small dataset. The MAE between prediction results and measurement results is about 0.02, which indicates that the predicted hysteresis Loops are very accurate reliable. The CNN U-net is not very sensitive to the learning rate and batch size, which means these two parameters can fluctuate over a relatively wide range and the CNN network can still function normally. This reduces the difficulty of the parameter-tunning. The comparison between the prediction results with the effectively trained baseline configuration CNN U-net and the measurement results are shown in Figure 8.

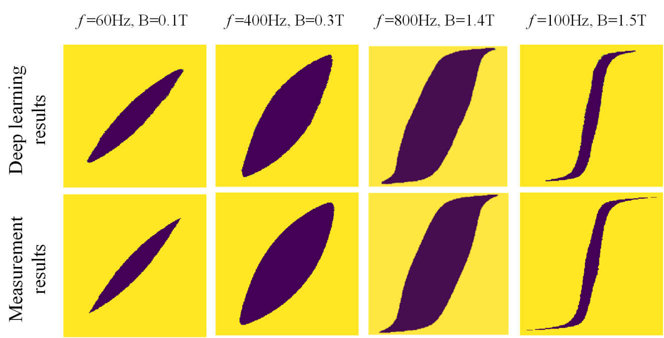


Figure 8. Comparison between predicted and measured hysteresis loops

For the 532 data sets of this experimental data, the proposed surrogate model can be effectively trained and converged with the only 40%–50% samples, which is about 213–266 samples. The trained model can accurately predict the remaining other samples, as shown in Figure. 8, the prediction results agree well with the measurements. This makes the model an effective tool to accelerate the measurement process of hysteresis loops as well as reduce the experimental cost and time.

4. CONCLUSION

In this paper, we have put forward a CNN U-net based approach to reduce the experimental cost and accelerate measurements for hysteresis loops of ferromagnetic materials. Take the measurement results obtained from the MPG200D Brockhaus equipment with an Epstein frame as ground truth, an extensive experimental analysis of supervised learning has been conducted. CNN U-net was adopted to explore the hidden relationship between physical parameters and B-H curves. Take parts of experimental measurement results as training samples for U-net, once the model can be effectively trained and converged, the rest of other hysteresis loops under different circumstances can be predicted without further measurement.

A set of empirical baseline hyperparameters configuration which can make the network converge effectively is put forward. With only 532 samples in total, the DL network can achieve effective learning after trained by barely 40%–50% of datasets. The prediction results agree well with the ground truth obtained from measurement. The trained model can output all the rest 256 prediction results in 1 minute. Our proposed method shows great potential in accelerating the measurement of the B-H curve. In addition, the sensitivity of CNN for hysteresis loops prediction with respect to the hyperparameters are investigated. Specifically, label normalization, training sample size, learning rate, and batch size are detailed discussed. This can guide future research of CNN for hysteresis loops prediction.

5. REFERENCES

[1] J. V. Leite, A. Benabou, and N. Sadowski, "Accurate minor loops calculation with a modified Jiles-Atherton hysteresis model," *Compel*, vol. 28, no. 3, pp. 741–749, 2009.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, pp. 1097–1105, Curran Associates, Inc., 2012.

[3] V. Zyuzin, P. Sergey, A. Mukhtarov, T. Chumarnaya, O. Solovyova, A. Bobkova, and V. Myasnikov, "Identification of the left ventricle endocardial border on two-dimensional ultrasound images using the convolutional neural network unet," in *2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, pp. 76–78, May 2018.

[4] X. Liu, W. Liu, T. Mei, and H. Ma, "A deep learning-based approach to progressive vehicle re-identification for urban surveillance," in *Computer Vision – ECCV 2016*, pp. 869–884, 2016.

[5] Z. Tang, X. Shen, and T. Henneron, "Application of U-net network and training strategy to optimal mesh refinement in computational electromagnetism," in *Conference on Electromagnetic Field Computation*, Oct. 2018.

[6] R. Gong and Z. Tang, "Investigation of convolutional neural network U-net under small datasets in transformer magneto-thermal coupled analysis," *Compel*, vol. 39, no. 4, pp. 959–970, 2020.

[7] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference for Learning Representations, San Diego*, 2014.

[8] A. Basu, S. De, A. Mukherjee, and E. Ullah, "Convergence guarantees for rmsprop and ADAM in non-convex optimization and their comparison to nesterov acceleration on autoencoders," *CoRR*, vol. abs/1807.06766, 2018.

[9] Y. Zhang and B. C. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification," *CoRR*, vol. abs/1510.03820, 2015.

[10] H. Sasaki and H. Igarashi, "Topology optimization accelerated by deep learning," *IEEE Transactions on Magnetics*, vol. 55, pp. 1–5, June 2019.