MagNet Challenge 2023 Final Report

Bowen Su, Yunhao Xiao, Min Yang, *Student Member, IEEE*, Kai Sun, *Fellow, IEEE*

***Abstract*- The power loss in magnetic materials when subjected to external changing magnetic flux is very important and can affects the efficiency of power equipment. However, most of the computing methods for losses in magnetic materials are still empirical fitting methods based on Steinmetz's equation. And the accuracy is affected by the Steinmetz's coefficients which vary with working conditions. A more accurate and flexible method is needed for the computation of losses in magnetic materials. Based on the database provided by the MagNet competition, this paper firstly introduces a loss prediction model using multilayer perceptron, achieving proper prediction results on different materials. Furthermore, to address the limitations of manually extracting input data features, a new prediction model for loss of magnetic materials is proposed based on the attention mechanism, further enhancing the accuracy of magnetic material loss predictions. The maximum and minimum relative errors for N87 are 24.717% and 0.003% respectively.**

I. Introduction

POWER losses are important performance parameters in the power equipment. It not only affects efficiency but is also closely related to the temperature of power equipment. Generally speaking, power losses can be divided into copper loss and iron loss. Copper loss refers to the power loss generated by the conductor when current flows through it. Currently, there have been many theoretical and computational methods available, such as the Dowell model and Tourkhani model [1], [2]. Iron loss refers to the power loss generated by magnetic materials when subjected to external changing magnetic flux. However, due to the lack of clear underlying physical models, there are rarely theories and computational models available for computing iron loss.

Currently, the main computation methods for iron loss are the Steinmetz's equation and its variations, including the Modified Steinmetz Equation, the Generalized Steinmetz Equation and Improved Steinmetz Equation [3]-[6]. These models are essentially based on the Steinmetz model, and their accuracy relies on the coefficients k, α, and β. However, the coefficients are very sensitive to operating conditions such as frequency and temperature, leading to significant errors in the existing methods. For example, the fitting coefficients α and β in the Steinmetz model will be affected by conditions such as different frequencies and magnetic induction intensities [7]. While current Steinmetz models and its variants takes the fitting coefficients α and β as constants. In addition, the above models hardly take into account the influence of temperature on the magnetic material losses. And the differences in losses at different tempe[i]ratures can be significant.

To address the above issues, there are two possible improvement approaches. On the one hand, a completely new model for calculating iron losses can be developed by analyzing and summarizing the characteristics of measured data sets. On the other hand, modern technologies such as artificial intelligence can be used to fit the Steinmetz coefficients using different methods, aiming to improve the accuracy of the formulas as much as possible. Numerous studies have been proposed in the field and are working towards the modeling of data-driven models for loss of magnetic material. A Machine Learning (ML)-based approach of identification for hysteresis and its inverse model are proposed in [8]. In [9], a neural network hysteresis model for the prediction of magnetic properties of ferromagnetic materials based on BP neural network is introduced, which only considered the influence of frequency on the magnetic properties. Recently, three neural network frameworks of the loss model of magnetic materials which cover different kinds of input and output, as well as a loss measurement system which can acquire large-scale and high-quality database are introduced in [10]. Furthermore, a transformer-based encoder-projector-decoder neural network architecture for modeling power magnetics B–H hysteresis loops is proposed, which can predict the losses of magnetic materials [11], [12]. However, magnetic material loss predictions based on hysteresis loops do not take into account the actual eddy current losses. In [13], a Deep Neural Network (DNN) approach to core loss estimations is proposed, which only considered the loss of N87 under a sinusoidal excitation.

Benefiting from the high-quality dataset provided by MagNet [14], this paper proposes a novel neural network structure for magnetic materials loss prediction, achieving high accuracy and model adaptability which can be used for different magnetic materials. In section II, the preliminary multi-layer perceptron regressor model for loss prediction of magnetic materials based on multiple input features extraction is proposed. Furthermore, an approach based on attention mechanism is proposed, which can address the limitations associated with the manual selection of features in the preliminary model. Section III gives the results of prediction errors for the testing materials.

II. Preahead Data Analysis and An Overview of Multi-layer Perceptron Model

A. Data analysis

Firstly, the data provided on April 1st, 2023 includes single cycle magnetic flux density waveforms (1024 discrete points) of ten different materials under various excitations, as well as

Bowen Su, Min Yang, Yunhao Xiao and Kai Sun are with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: subw20@mails.tsinghua.edu.cn,; yangmin20@mails.tsinghua.edu.cn; xiaoyh23@mails.tsinghua.edu.cn; sun-kai@mail.tsinghua.edu.cn).

the corresponding excitation waveform frequencies, operating temperatures, and single cycle magnetic field intensity waveforms (1024 discrete points).

After analysis, the provided single cycle magnetic flux density waveforms have significant differences, not only in the standard sine wave waveform, but also in a variety of magnetic flux density waveforms with different harmonic contents, as shown in Fig. 1.

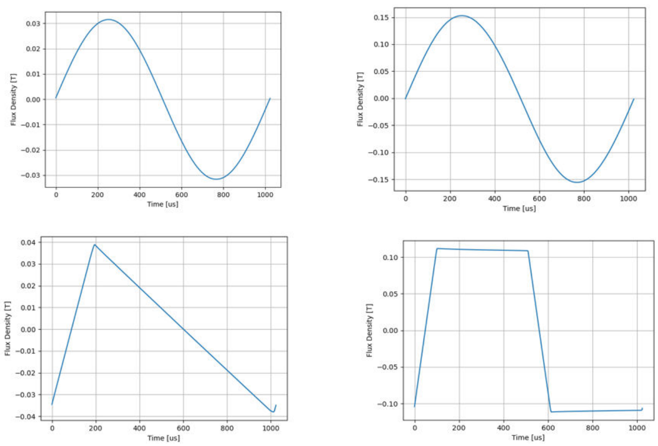

Fig.1 Waveforms of different magnetic flux density

Secondly, the data provided for final evaluation on November 10th, 2023 includes single cycle magnetic flux density waveforms (1024 discrete points) of 5 different materials under different excitation waveforms, as well as the corresponding frequency and operating temperature of the excitation waveforms. And consistent with the data provided for the first time, the provided single cycle magnetic flux density waveform still has significant differences, containing multiple magnetic flux density waveforms with different harmonic contents, as shown in Fig. 2.

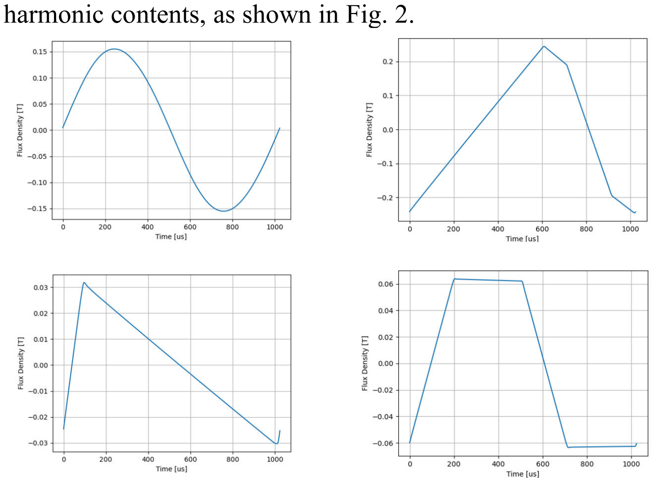

Fig.2 Waveforms of different magnetic flux density in final evaluation

However, it can be observed that the waveform shape of the data provided for testing is roughly the same as the waveform provided for the first time. Therefore, the deep learning model trained from the waveform provided for the first time can be well transferred to the waveform used for testing, generating robust magnetic core loss prediction results.

B. Multi-layer perceptron regressor model

According to the data analysis in the previous text, it can be seen that the given data mainly consists of magnetic flux density waveforms and related working conditions, and the magnetic flux density waveforms are relatively complex, containing various harmonics with different frequencies and amplitudes.

Therefore, the calculation of various features that comprehensively reflect the waveform properties of magnetic flux density is envisioned, followed by the determination of magnetic core loss density through the utilization of a Multi-layer Perceptron (MLP) regressor.

Concerning feature selection in the time-domain, the amplitude, peak-to-peak value, effective value, and variance of the magnetic flux density waveform are directly extracted as four dimensional features. Subsequently, two types of dimensionless features, namely skewness and kurtosis, of the magnetic flux density waveform are computed. The calculation formula is shown in (1), (2).

$$S = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma_i}\right)^3 \tag{1}$$

$$K = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma_i}\right)^4 \tag{2}$$

Regarding frequency domain characteristics, Fast Fourier Transform (FFT) is initially applied to the magnetic flux density waveform. Essential frequency domain features, including fundamental amplitude and frequency, second harmonic amplitude and frequency, as well as third harmonic amplitude and third harmonic frequency, are then extracted from the spectrum of the magnetic flux density waveform.

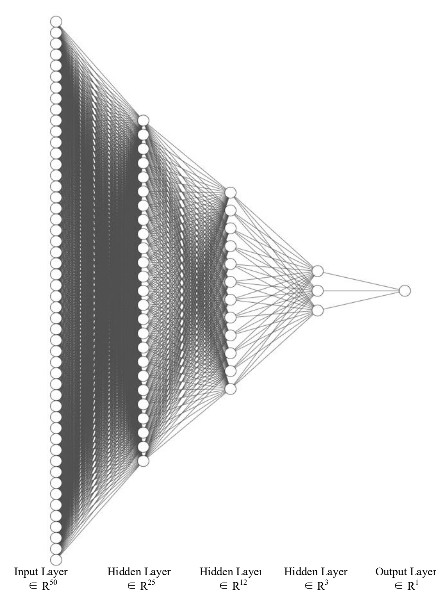

Fig.3 The structure of MLP

To effectively capture potential coupling relationships among the selected features, polynomial features of degree 3 or higher are incorporated for all features. Utilizing the product method, the relationships between the aforementioned features are indirectly characterized within the newly introduced polynomial features.

The MLP regressor consists of three fully connected layers, with each layer corresponding to a number of neurons (210, 150, 30). And after multiple tests, the GELU function is chosen as the activation function, which has higher accuracy and is more conducive to the back propagation process. The model structure is shown in Fig. 3.

However, this method of manually selecting features and applying MLP regressors for core loss prediction has two drawbacks. The first issue is that the model parameters are too large. In order to better preserve waveform information, more features will be selected, which leads to an increase in the number of neurons in the fully connected layer, thereby greatly increasing the overall parameters of the model; The second reason is that the model accuracy is relatively low, as the manual feature selection process is greatly influenced by the operator's professional knowledge and experience, and the features obtained from different magnetic flux density waveforms through the same feature selection process may vary greatly, resulting in lower robustness.

III. The Proposed Encoder-based Model

In order to overcome the drawbacks of manually selecting features and applying MLP regressor methods mentioned above, an attention mechanism based approach from sequence to scale is proposed, with the structure shown in Fig. 4.

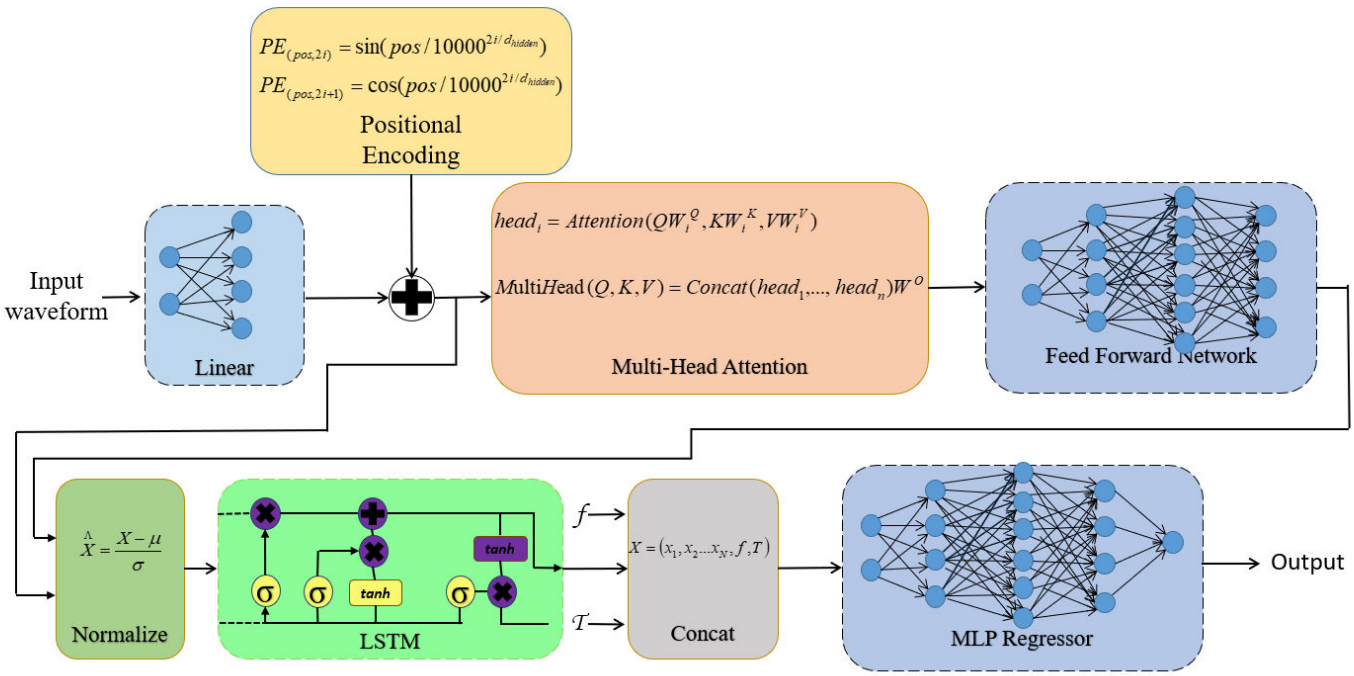

Fig.4 The structure of encoder-based model

In this method, the magnetic flux density waveform with 1024 discrete points is first directly taken as input, and each discrete single point value is expanded into a 64-dimensional vector. After expansion, more information may be stored during the training process, which is more conducive to subsequent attention calculation. Then, its position is encoded by (3), (4).

$$PE_{(pos,2i)} = \sin(pos/10000^{2i/d_{hidden}}) \quad (3)$$

$$PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d_{hidden}}) \quad (4)$$

The use of sine function as position encoding mainly takes into account that due to the unique mathematical manipulation property of the sine function, any dimensional position encoding of any point can be linearly represented by the position encoding of other points, thus effectively characterizing the relative position relationship between points.

Afterwards, the waveform with added position encoding will be input into an encoder layer[15] with self-attention mechanism. In this layer, the internal relationships of the magnetic flux density waveform will be constructed through self-attention mechanism, and the relationships between different points will be fully constructed through different attention scores, which is calculated as the formula (5). After passing through the encoder layer, each vector contains information about the entire waveform and has different values depending on the attention score. And the whole process of the attention mechanism is shown in Fig.5.

$$Attention(Q,K,V) = Softmax(\frac{QK^T}{\sqrt{d_k}}) \quad (5)$$

The attention mechanism currently lacks a comprehensive and intuitive mathematical explanation, but intuitively speaking, it models the strength of the interrelationships between different input vectors by extracting the Query matrix and Key matrix of the input matrix and performing matrix dot multiplication on them. Intuitively speaking, the attention mechanism allows the model to see the key information and possible connection relationships in the input data, effectively modeling the internal relationships of the input waveform. In this case, after the input waveform matrix passes through an encoder layer containing self attention, each vector in the matrix contains the overall information of the matrix, which also provides the possibility for effective feature extraction in subsequent structures.

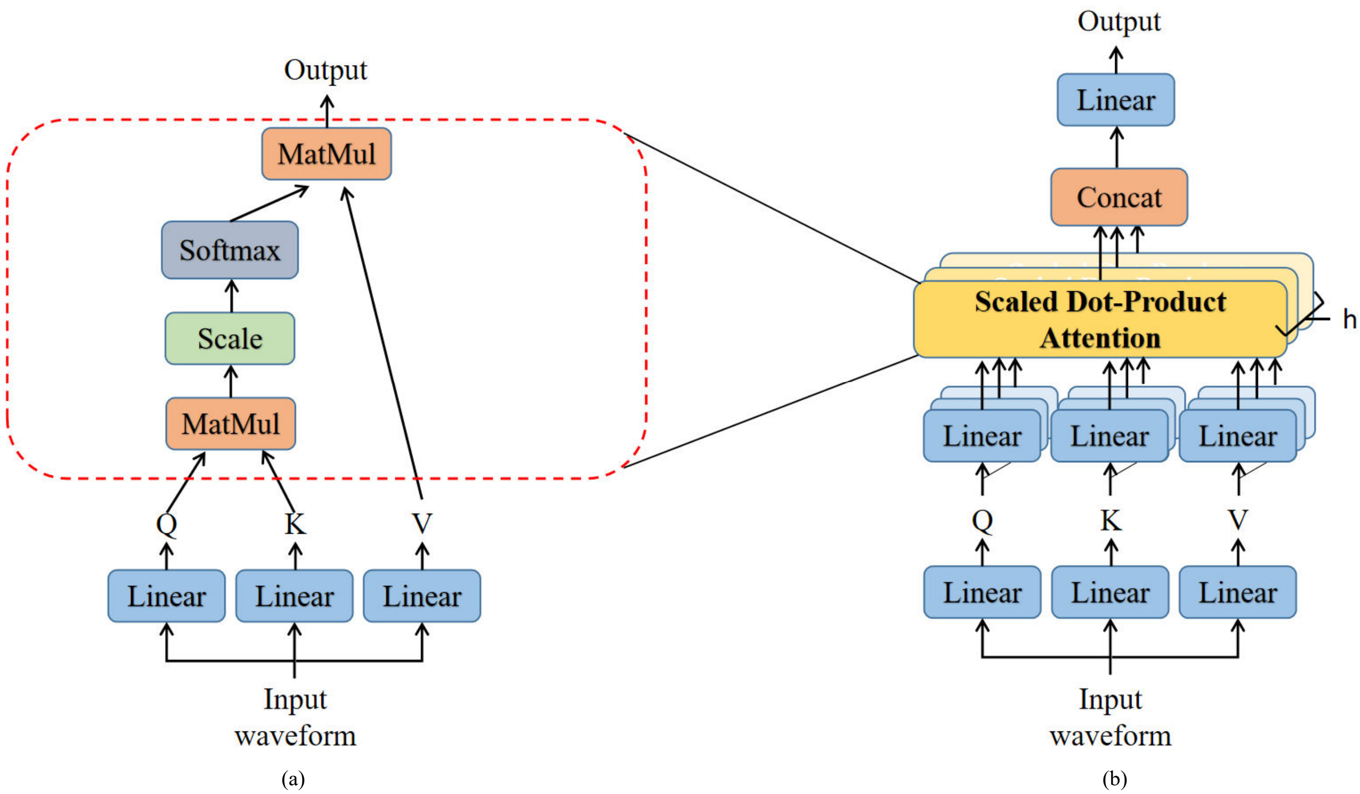

Fig.5 The process of attention mechanism.(a) describes self-attention mechanism and (b) describes the multi-head attention mechanism.

Then, feature extraction is performed using a Long Short-Term Memory (LSTM)[16] network. A matrix containing 1024 vectors is input into the LSTM layer, and each vector of the matrix is treated as a sequence of information. Through continuous iteration of the LSTM, the last intermediate state is finally taken as the enrichment of the entire matrix information.

Finally, the extracted information vector is partially compressed through a simple fully connected network, then concatenated with the excitation waveform frequency and operating temperature, and input as input features into the MLP regressor. A simple double-layer fully connected network with a structure of (16,8) is used to predict the magnetic core loss density.

Overall, the proposed model mainly utilizes the encoder layer and its inherent self-attention mechanism to expand the information and model the internal relationships of the magnetic flux density waveform, then uses LSTM for feature extraction, and finally uses MLP regression to predict the magnetic core loss density. In fact, the adoption of self-attention mechanism plays an excellent role in modeling the internal relationships of waveforms, and this mechanism does not bring excessive parameter increase or other negative effects to the entire model. Moreover, since the self-attention mechanism is independent of parameters, it can be applied to different input waveforms, even if these waveforms may be very complex and very different from those in the training data, which undoubtedly greatly increases the robustness of the model and effectively improves the accuracy of the model.

Although the above model has good performance in accuracy and robustness, it also has a significant drawback, which is its high time complexity. Because time-series processing models such as Encoder layer and LSTM are used in the model, it means that the complexity of the model is closely related to the length of the processed sequence. For the self-attention mechanism, for a sequence of length N, its complexity is O (N^2), while for LSTM, for a sequence of length N, its complexity is O (N). Therefore, when the sequence length is long, that is, when the sampling rate of the single cycle magnetic flux density waveform is high, the computational complexity of the model will significantly increase. Correspondingly, the training time and deployed running time will also increase significantly.

In order to further improve accuracy, some hyperparameters can be adjusted and optimized, such as the number of layers in the encoder layer in this model. Using multi-layer encoders can more accurately model the internal relationship of the magnetic flux density waveform, but it should be noted that this will increase the model parameters and computational complexity. In addition, the feature extraction part of the model can be modified. Except for LSTM, structures such as MLP or convolutional neural networks can also play a role in feature extraction. In summary, there is still room for further improvement in this model.

IV. RESULTS

The training process for the five materials used for final evaluation is basically similar. Firstly, the given training data for each material is divided into a training set and a validation set in a 7:3 ratio, and the hyperparameters are adjusted on the validation set.

And the Adam optimizer[17] is used with β_1 = 0.9, β_2 = 0.999 and ξ = 10^{-9}. The decay of the learning rate over the course of training is as shown in (6).

$$lrate = lrate_{initial} \times (0.5)^{epoch \| 100} \tag{6}$$

This corresponds to decreasing the learning rate as the number of training epoch increases, which is beneficial for the model to converge to its optimal point. The number of parameters of each model for different material is shown in TABLE I.

TABLE I Number of parameters in models for different materials

Materials	A	B	C	D	E
Parameter-number	116061	116061	116061	116061	116061

And the test results of the two methods mentioned earlier, namely MLP regressor model and Encoder based model, on material N87, are shown in Table Ⅱ.

TABLE II Test results of different models for N87 material

Method	Min. error	Avg. error	Max. error	95th. error	R2 Score
MLP	0.002%	13.642%	101.964%	23.336%	0.974
LSTM	4.791%	3.668%	26.185%	9.439%	0.997
Proposed model	0.003%	3.568%	24.717%	8.707%	0.998

V. CONCLUSION

This paper utilizes the dataset provided by MagNet. Firstly, a magnetic material loss prediction model is proposed based on a feedforward neural network, focusing on the extraction of magnetic flux density waveform features. The testing errors for different materials are maintained within 10%. Considering the limitations of manually extracting waveform features, this paper further proposes an encoder-based model which contains attention mechanism to address the limitations associated with the manual selection of features. Compared with the MPL model and LSTM model, the proposed encoder-based model demonstrates higher accuracy in predicting the losses of magnetic materials.

In the future, there are still many potential directions for the development of proposed model. On the one hand, incorporating some physical mechanisms into the consideration of neural network models can enhance the interpretability of the neural network at the physical level. On the other hand, further adjustments to the network structure can be made to extract input features more effectively, thereby improving the model's accuracy.

REFERENCES

[1] P. L. Dowell, "Effects of eddy current in transformer windings", *Proc. I nst. Electr. Eng.*, vol. 113, no. 8, pp. 1387-1394, Aug. 1966.

[2] F. Tourkhani and P. Viarouge, "Accurate analytical model of winding lo sses in round Litz wire windings," in *IEEE Transactions on Magnetics*, vol. 37, no. 1, pp. 538-543, Jan 2001.

[3] C. P. Steinmetz, "On the law of hysteresis," in *Proceedings of the IEEE*, vol. 72, no. 2, pp. 197-221, Feb. 1984.

[4] J. Reinert, A. Brockmeyer and R. W. A. A. De Doncker, "Calculation of losses in ferro- and ferrimagnetic materials based on the modified Stein metz equation," in *IEEE Transactions on Industry Applications*, vol. 37, no. 4, pp. 1055-1061, July-Aug. 2001.

[5] Jieli Li, T. Abdallah and C. R. Sullivan, "Improved calculation of core l oss with nonsinusoidal waveforms," *Conference Record of the 2001 IEE E Industry Applications Conference. 36th IAS Annual Meeting (Cat. No. 01CH37248)*, Chicago, IL, USA, 2001, pp. 2203-2210 vol.4.

[6] K. Venkatachalam, C. R. Sullivan, T. Abdallah and H. Tacca, "Accurate prediction of ferrite core loss with nonsinusoidal waveforms using only Steinmetz parameters," *2002 IEEE Workshop on Computers in Power E lectronics, 2002. Proceedings.*, Mayaguez, PR, USA, 2002, pp. 36-41.

[7] P. Papamanolis, T. Guillod, F. Krismer and J. W. Kolar, "Minimum Los s Operation and Optimal Design of High-Frequency Inductors for Defin ed Core and Litz Wire," in *IEEE Open Journal of Power Electronics*, vo l. 1, pp. 469-487, 2020.

[8] X. Zhao and Y. Tan, "Modeling Hysteresis and Its Inverse Model Using Neural Networks Based on Expanded Input Space Method," in *IEEE Tr ansactions on Control Systems Technology*, vol. 16, no. 3, pp. 484-490, May 2008.

[9] M. Tian, H. Li and H. Zhang, "Neural Network Model for Magnetizatio n Characteristics of Ferromagnetic Materials," in *IEEE Access*, vol. 9, p p. 71236-71243, 2021.

[10] H. Li *et al*., "How MagNet: Machine Learning Framework for Modeling Power Magnetic Material Characteristics," in *IEEE Transactions on Po wer Electronics*, vol. 38, no. 12, pp. 15829-15853, Dec. 2023.

[11] H. Li, D. Serrano, S. Wang, T. Guillod, M. Luo and M. Chen, "Predictin g the B-H Loops of Power Magnetics with Transformer-based Encoder-Projector-Decoder Neural Network Architecture," *2023 IEEE Applied P ower Electronics Conference and Exposition (APEC)*, Orlando, FL, US

A, 2023, pp. 1543-1550.

[12] H. Li, D. Serrano, S. Wang and M. Chen, "MagNet-AI: Neural Network as Datasheet for Magnetics Modeling and Material Recommendation," in *IEEE Transactions on Power Electronics*, vol. 38, no. 12, pp. 15854-15869, Dec. 2023.

[13] X. Shen, H. Wouters and W. Martinez, "Deep Neural Network for Magnetic Core Loss Estimation using the MagNet Experimental Database," *2022 24th European Conference on Power Electronics and Applications (EPE'22 ECCE Europe)*, Hanover, Germany, 2022, pp. 1-8.

[14] H. Li et al., "MagNet: An Open-Source Database for Data-Driven Magnetic Core Loss Modeling," *2022 IEEE Applied Power Electronics Conference and Exposition (APEC)*, Houston, TX, USA, 2022, pp. 588-595.

[15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, "Attention Is All You Need," arXiv:1706.03762 [cs.CL], Jun, 2017. DOI: 10.48550/arxiv.1706.03762.

[16] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997.

[17] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization," arXiv:1412.6980 [cs.LG], Dec, 2014. DOI: 10.48550/arXiv.1412.6980