MagNet Challenge 2023-Ferrite Core Loss Model Based on LSTM and Transfer Learning

Zhanlei Liu, Cao Zhan, Yongliang Dang, Yukun Zhang, Na Wang, Yiting Chen, Yiming Zhang

***Abstract*- The accurate modelling of core loss is significant for the temperature prediction and thermal design of magnetic components. Existing core loss calculation methods for ferrite are either inaccurate or time-consuming, hindering the precise design of magnetic components. In this paper, a neural network based ferrite core loss model is proposed. The CNN+LSTM network is proposed the model the core loss on large datasets. To reduce the demand for data size, a transfer learning network is utilized to model the core loss on small dataset. The proposed LSTM and transfer learning based network can achieve accurate core loss calculation.**

I. INTRODUCTION

High frequency transformers (HFT) and high frequency filter inductors (HFFI) are widely used in DC-DC converters such as grid-connected inverters, railway transportations and power supply of data centers. Magnetic components are key components in HFTs and HFFIs, which provide the roles of power transmission and energy storage. The modelling and design of magnetic components is essential in the design of HFTs and HFFIs.

The core loss can account for half of the total power loss in magnetic components. Accurate modelling of core loss is one of the key steps in the modelling of magnetic components. Various core loss calculation methods had been proposed. The Original Steinmetz equation (OSE) proposed by Steinmetz is simple and practical, which is limited to sinusoidal excitations. In order to extend the empirical model to high frequency and non-sinusoidal excitations, some modified Steinmetz equations were proposed, including the MSE, IGSE, WcSE, I^2GSE, and CWH. However, the accuracy of Steinmetz equations in wide ranges of frequency and flux density is low. The loss separation method proposed by Bertotti cannot achieve accurate core loss calculation for ferrite under high frequency excitations. The hysteresis models are too complex to apply in practical engineering. The neural network is a data-driven model, which not relies on the specific core loss mechanisms. The researchers in Princeton and Dartmouth have collected large ferrite core loss data and established some neural network models to predict the ferrite core loss such as feedforward neural network, transfer learning and long short-term memory network models. The models can predict the core loss under arbitrary waveform excitations and the predicted core loss showed higher accuracy than traditional empirical equations. However, the model complexities are high and the core loss mechanisms are not explained in the models.

In this paper, a neural network based ferrite core loss model is proposed. The CNN+LSTM network is proposed to model the core loss under large core loss datasets. The transfer learning method is utilized to model the core loss under small core loss datasets. The accuracy of these neural network models are analyzed.

II. CNN+LSTM NETWORK BASED CORE LOSS MODEL ON LARGE DATASET

A. Analysis of the Large Core Loss Dataset

The core loss dataset used in this model is established by the researchers from Dartmouth and Princeton Universities, including various materials (3C90, 3C94, 3E6, and et al.), wide frequency (50kHz-500kHz), wide flux density (10mT-300mT), wide temperature (25℃-90℃) and several excitation waveforms (sinusoidal, triangle, and trapezoidal). The datasets include flux density sequences, frequencies, temperatures and core loss densities.

B. CNN+LSTM Network

A CNN+LSTM network is proposed to model the core loss as shown in Fig. 1. The model takes the flux density sequence, temperature, and frequency as input and the core loss density as output. The CNN is used to extract the local features in the flux density sequence and the LSTM is used to extract the time series features in the flux density sequence. The FCs are used to characterize the core loss density with flux density sequence features, frequency and temperature.

In this model, the input layer has a feature dimension of 16. The CNN module consists of a convolutional layer, ReLU activation function, and max pooling layer, using 1D convolution for feature extraction. It has 1 input channel, 16 output channels, a convolution kernel size of 25, and a pooling window size of 10. This module is used to extract and integrate features from multiple batches of time series data. The LSTM layer models the input time series data to extract higher-level feature representations. The use of Bi-LSTM helps capture past and future contexts of input elements, maintaining long-term dependency relationships between model parameters and learning data. Additionally, it helps handle variable-length sequences and can process sequences of different lengths in batches. The input size of the feature dimension after convolutional kernel pooling is 16. The hidden size determines the output dimension of the LSTM and increases model complexity and training time. A larger hidden size can provide richer feature expression capability. The hidden size of this model is 32. The number of layers determines the depth of the LSTM network, allowing for more complex sequence patterns. This model has 1 layer. The final output size of the Bi-LSTM is 64. After the first fully connected layer with the activation function ReLU, the output size is 32. The magnetic field intensity features extracted from the previous steps are then combined with peak intensity, frequency, and temperature to form a sequence of size 36. After the second fully connected layer with the activation function ReLU, the output size is 16. Finally, the loss value is obtained through the last fully connected layer. Dropout controls the regularization level of the model, and a larger dropout can reduce overfitting to some extent and reduce the training burden of the model. In addition, the batch size is 32, the number of epochs is 200, and the learning rate is 0.01. The CNN has a total of 5068 parameters, including the convolutional and pooling layers, while the Bi-LSTM has a total of 54784 parameters, including the hidden and fully connected layers.

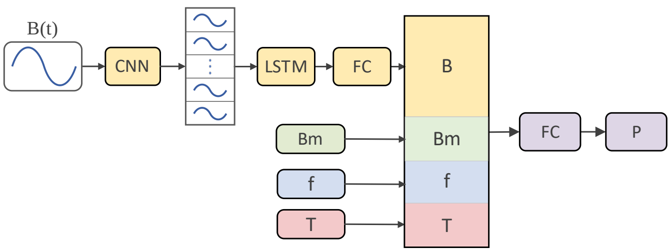
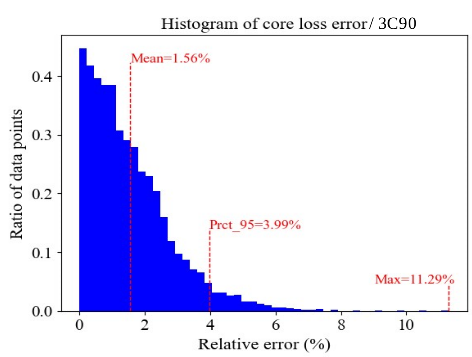


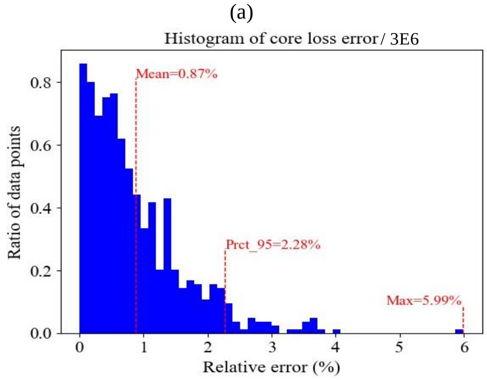
Fig. 1 Histograms of core loss error under large datasets. (a) 3C90, (b) 3E6, (c) 77, (d) N27.

C. Results

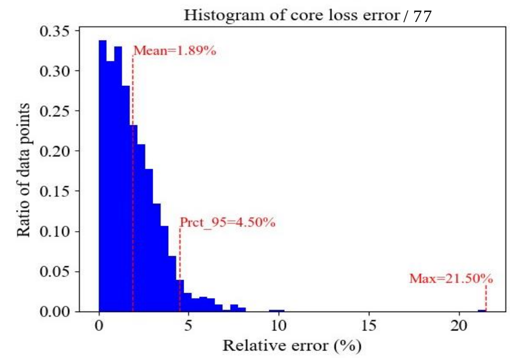
The errors of the proposed core loss model are shown in Fig. 2. The mean errors are within 3%, while error of the top 95% of the data are within 5%, Particularly for 3E6, the mean error is about 0.87%, while the error of the top 95% of the data is below 2.5%.



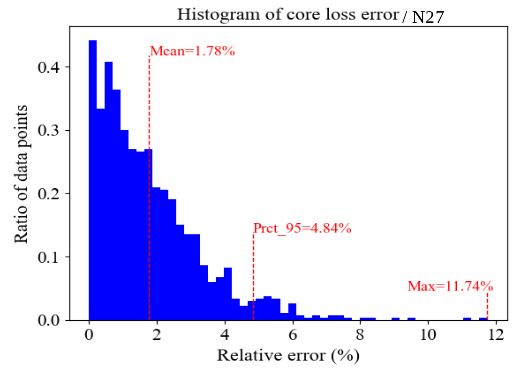
(a)



(b)



(c)



(d)

Fig. 2 Histograms of core loss error based on large datasets. (a) 3C90, (b) 3E6, (c) 77, (d) N27.

III. Transfer Learning Based Core Loss Model on Small Dataset

The neural network model based on CNN+LSTM can realize accurate modeling and calculation of ferrite core loss. However, the above model has high requirements on the number of samples in the training set and needs to be modeled based on loss data stimulated by wide frequency, wide magnetic flux density, wide temperature and various complex waveforms, which is difficult to be applied in practical engineering. Therefore, how to use small sample data set to achieve core loss modeling is of great significance for practical engineering.

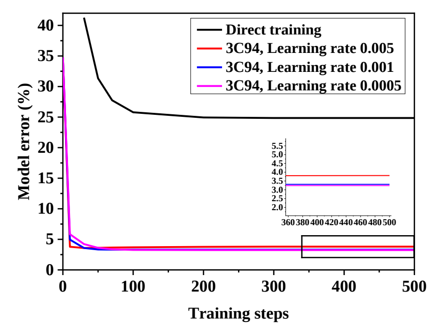
Transfer learning (TL) is a new machine learning method developed in recent years. It uses the source task to pre-train the model, and then fine-tunes the pre-trained model through the target task to obtain the learning model of the target task. It can be used to solve the problem of insufficient data. Compared with traditional machine learning, it is more efficient and accurate. The TL model based on parameter fine-tuning is simple and fast in calculation, and is a common method in TL. The fine-tuning method first trains a pre-trained model on a large sample data set, then freezes the network parameters of the first few layers of the pre-trained model, and trains and fine-tunes the basic network parameters behind the pre-trained model based on the new small sample data. Finally, the pre-trained model is migrated to new data samples.

TABLE I
THE PRE-TRAINED MODEL IS DIRECTLY APPLIED TO THE PREDICTION ACCURACY OF 3C90 IN A LARGE SAMPLE DATASET

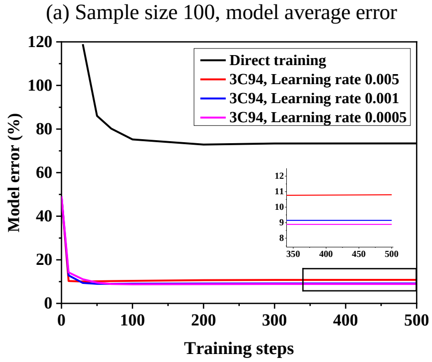
Pretraining model	3C94	3E6	3F4	77	78	N27	N30	N49	N87
Model mean error /%	11.1	165.8	113.74	19.79	15.05	32.44	137.36	52.89	48.8
Model 95% error /%	28.54	461.01	324.3	35.43	29.52	52.07	331.11	143.41	67.03

In order to make the small sample dataset of five new materials achieve better prediction effect, this paper established a ferrite loss model based on TL on the basis of the large sample data set neural network model and the similar mechanism between different types of ferrite core losses. In order to verify the validity of the parametric fine-tuning loss prediction model, this paper takes the 3C90 model ferrite as an example to construct 5 small sample datasets, the sample sizes of which are 100, 200, 400, 600 and 1000 respectively. The test set samples are the remaining samples after the total data set samples are removed from the training set samples. Then evaluate the prediction accuracy of nine other models of ferrite core loss pre-training models directly applied to 3C90, as shown in TABLE I below. It can be seen that 3C94 ferrite core loss pre-training model has the highest accuracy when applied to 3C90. Therefore, the training results of 3C94 ferrite core loss model will be used as the pre-training model below, and 3C90 small sample data set will be used to fine-tune the parameters of the pre-training model to achieve model migration.

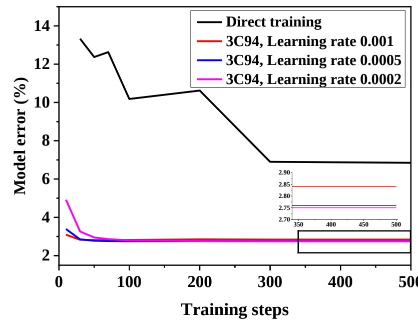
For the 3C90 ferrite core loss small sample data set, the model errors of direct training and TL training were compared, as shown in Fig. 3. It can be seen that for small sample data sets, TL can greatly improve the convergence speed and accuracy of the model, and better model performance can be obtained by selecting an appropriate learning rate. In addition, with the increase of the number of training samples, the accuracy of the model is improved.



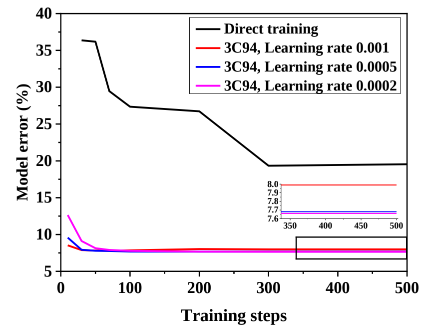
(a) Sample size 100, model average error



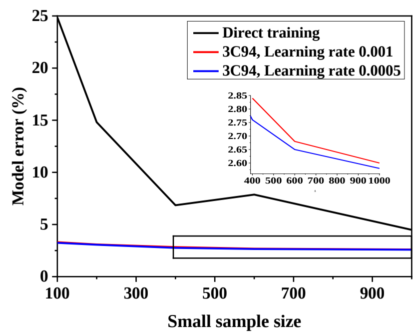
(b) Sample size 100, 95% error of the model



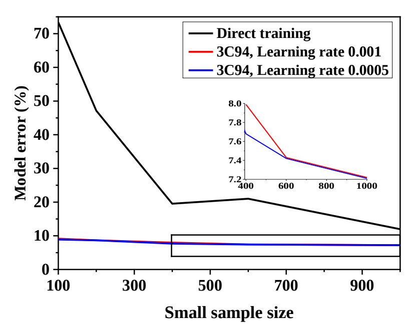
(c) Sample size 400, model average error



(d) Sample size 400, 95% error of the model



(e) Model average error



(f) 95% error of the model

Fig. 3 Comparison of model errors between direct training and TL training in small sample dataset.

Therefore, in order to make the loss prediction of the 5 unknown new materials more accurate, the prediction accuracy of other 10 known models of ferrite core loss pre-training model directly applied to the 5 new materials was evaluated first, and the training results of the known model of ferrite loss with the highest prediction accuracy were selected as the pre-training model of the unknown new materials. Finally, the sample data set of the unknown new material was used to fine-tune the parameters of the pre-trained model to achieve model migration and accurate prediction of the loss of the new material with small sample data set. The specific process is shown in Fig. 4.

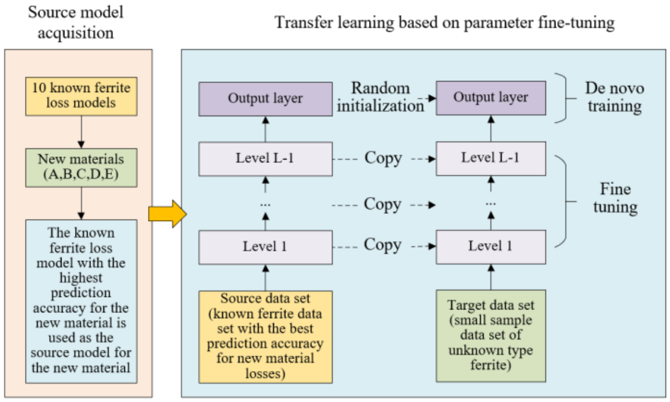


Fig. 4 Prediction of ferrite loss of unknown type based on parametric fine-tuning TL

In transfer learning, the input layer has a feature dimension size of 16. Firstly, a CNN module consisting of a convolutional layer, ReLU activation function, and max pooling layer is used for feature extraction. The 1D convolution performs feature extraction with 1 input channel, 16 output channels, a kernel size of 25, a pooling window size of 10, and a stride of 10 to extract and integrate multiple batches of time series features. The LSTM layer contains a bidirectional LSTM layer with an input size of 16, a hidden size of 32, and a layer number of 1. The output of the LSTM layer is used as the input to the fully connected layer. Fully connected layer 1 consists of three linear layers with a structure of 32-32-13 and ReLU activation function. Fully connected layer 2 consists of three linear layers with a structure of 16-16-1 and ReLU activation function. The optimizer selected is Adam optimizer with a learning rate of 0.0001. The special feature of the optimizer is that it uses the filter function to select the parameters with requires 'grad=True' in the model for optimization. The total number of parameters in the model is 7154.

IV. Conclusion

This paper proposed a neural network based core loss model for ferrite. The CNN+LSTM network is proposed to model the core loss under large core loss datasets. The transfer learning method is utilized to model the core loss under small core loss datasets. The accuracy of these neural network models are high.

References

[1] C. P. Steinmetz, "On the law of hysteresis," *Proc. IEEE*, vol. 72, no. 2, pp. 197–221, Feb. 1984.

[2] J. Reinert, A. Brockmeyer, and R. W. A. A. De Doncker, "Calculation of losses in ferro- and ferrimagnetic materials based on the modified Steinmetz equation," *IEEE Trans. Ind. Appl.*, vol. 37, no. 4, pp. 1055–1061, Jul./Aug. 2001.

[3] K. Venkatachalam, C. R. Sullivan, T. Abdallah, and H. Tacca, "Accurate prediction of ferrite core loss with nonsinusoidal waveforms using only Steinmetz parameters," *Proc. IEEE Workshop Comput. Power Electron.*, Jun. 2002, pp. 36–41.

[4] W. Shen, F. Wang, D. Boroyevich, and C. W. Tipton, "Loss characterization and calculation of nanocrystalline cores for high-frequency magnetics applications," *IEEE Trans. Power Electron.*, vol. 23, no. 1, pp. 475–484, Jan. 2008.

[5] J. Muhlethaler, J. Biela, J. W. Kolar, and A. Ecklebe, "Improved core loss calculation for magnetic components employed in power electronic systems," *IEEE Trans. Power Electron.*, vol. 27, no. 2, pp. 964–973, Feb. 2012.

[6] M. Mu, and F. C. Lee, "A new core loss model for rectangular AC voltages," in *Proc. Energy Convers. Congr. Expo. Conf.*, 2014, pp. 5214– 5220.

[7] C. R. Sullivan, J. H. Harris, and E. Herbert, "Core loss predictions for general PWM waveforms from a simplified set of measured data," in *Proc. 25th Annu. IEEE Appl. Power Electron. Conf. Expo.*, 2010, pp. 1048–1055.

[8] S. Barg, K. Ammous, H. Mejbri, and A. Ammous, "An improved empirical formulation for magnetic core losses estimation under nonsinusoidal induction," *IEEE Trans. Power Electron.*, vol. 32, no. 3, pp. 2146–2154, Mar. 2017.

[9] G. Bertotti, "General properties of power losses in soft ferromagnetic materials," *IEEE Trans. Magn.*, vol. MAG-24, no. 1, pp. 621–630, Jan. 1988.

[10] D. C. Jiles and D. L. Atherton, "Theory of ferromagnetic hysteresis," *J. Magn. Magn. Mater.*, vol. 61, pp. 48–60, 1986.

[11] Y. Bernard, E. Mendes, and F. Bouillault, "Dynamic hysteresis modeling based on Preisach model," *IEEE Trans. Magn.*, vol. 38, no. 2, pp. 885–888, Mar. 2002.

[12] E. Dogariu, H. Li, S. Wang, M. Luo and M. Chen, "Transfer learning methods for data-driven magnetic core loss modeling," *IEEE Workshop on Control and Modeling of Power Electronics (COMPEL)*, 2021.

[13] H. Li, S. R. Lee, M. Luo, C. R. Sullivan, Y. Chen and M. Chen, "MagNet: A Machine Learning Framework for Magnetic Core Loss Modeling," in *Proc. IEEE 21st Workshop Control Model. Power Electron.*. 2020, pp. 1–8.

[14] H. Li et al., "MagNet: An Open-Source Database for Data-Driven Magnetic Core Loss Modeling," in *Proc. IEEE Appl. Power Electron. Conf. Expo.*, 2022, pp. 588-595.

[15] D. Serrano et al., "Neural Network as Datasheet: Modeling B-H Loops of Power Magnetics with Sequence-to-Sequence LSTM Encoder-Decoder Architecture," in *Proc. IEEE 23rd Workshop Control Model. Power Electron.*, 2022, pp. 1–8.

[16] E. Dogariu, H. Li, D. Serrano López, S. Wang, M. Luo, and M. Chen, "Transfer Learning Methods for Magnetic Core Loss Modeling," in *Proc. IEEE 22nd Workshop Control Model. Power Electron.*, 2021, pp. 1-6.

[17] D. Serrano et al., "Why MagNet: Quantifying the Complexity of Modeling Power Magnetic Material Characteristics," *IEEE Trans. Power Electron.*, vol. 38, no. 11, pp. 14292-14316, Nov. 2023.

[18] H. Li et al., "How MagNet: Machine Learning Framework for Modeling Power Magnetic Material Characteristics," *IEEE Trans. Power Electron.*, vol. 38, no. 12, pp. 15829-15853, Dec. 2023.

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