MagNet Challenge 2023 Final Report

[*]Yun-Shan Hsieh, [*]Jian-De Li, [*]Li-Chen Yu, [*]Tzu-Chieh Hsu, [*]Yu-Chen, Liu, *Senior Member, IEEE,*
[+]Chin-Hsien Hsia, *Senior Member, IEEE,* [§] Chen Chen, *Member, IEEE*

[*]*National Taipei University of Technology, Taiwan*
[+]*National Ilan University, Taiwan*
[§]*LITEON Technology, Taiwan*

Abstract- In this report, a novel architecture of a Feedforward Neural Network (FNN) is proposed for the prediction of core loss. The architecture was refined using Optuna, resulting in a network configuration with five hidden layers, consisting of 65, 55, 116, 40, and 123 neurons, respectively. The model's hyperparameters were optimized, featuring a decay epoch of 423, a decay ratio of 0.458, and an initial learning rate set at 0.005. To enhance the model's precision and operational efficiency, Z-score normalization techniques were employed. Addressing the challenge of scarce training data, the study integrated transfer learning with the Dynamic Time Warping (DTW) algorithm. This integration facilitated the evaluation of the correlation between 10 pre-test materials and an unknown material, leading to the selection of the most pertinent materials for pre-training. It was observed that pre-training did not significantly improve the results for Materials A and B, thereby restricting the application of transfer learning to Materials C, D, and E. The dataset was divided into 60% for training, 20% for validation, and 20% for testing purposes. The results indicated that all materials, with the exception of Material D, achieved an average error rate below 5%.

I. Introduction

Magnetic components in converters, which often operate under high-frequency conditions, are subject to substantial variations in their excitation sources. These sources can encompass a range of waveforms, including square waves, triangular waves, and those incorporating a DC component. Traditionally, the calculation of magnetic core losses in such environments has predominantly relied on the Steinmetz Equation (SE) or the Original Steinmetz Equation (OSE) [1]. These formulas are primarily designed for sine wave excitations that do not involve a DC bias, as depicted in Equation (1). This conventional approach, while effective under specific conditions, may not fully encapsulate the complexities encountered in varied waveform scenarios.

$$P_{SE} = kf^{\alpha}B_S^{\beta} \tag{1}$$

However, in switch-mode power supplies, sine waves are typically not used, and the excitation source often appears in the form of triangular waves or square waves. Different excitation sources generate different magnetic hysteresis loops. If the same loss calculation method is still used, significant errors will be introduced. Therefore, from the SE, various calculation equations have been derived to accommodate non-sinusoidal excitations and suit different scenarios. These include the Generalized Steinmetz Equation (GSE) [2], Improved Generalized Steinmetz Equation (iGSE) [3], I2 Generalized Steinmetz Equation (i^{2}GSE) [4], and Equivalent Elliptical Loop (EEL) [5]. Unfortunately, these loss calculation methods are not able to fully capture the differences in different temperatures or waveforms with a bias current. Therefore, there is still room for improvement in the current loss equations.

The emergence of neural networks has ushered in a transformative phase in machine learning, with deep neural networks (DNNs) standing out as a particularly revolutionary advancement. Exhibiting exceptional aptitude for tackling complex, nonlinear issues, which are commonplace in multivariate classification and regression, these networks are distinguished by their sophisticated architectures and an innate ability to identify intricate patterns. This aptitude renders DNNs exceptionally suitable for challenging tasks such as core loss prediction, characterized by its nonlinear and multi-dimensional nature. Our report is dedicated to elaborating on the utilization of deep neural networks for the precise prediction of core losses.

Section 2 delves into the detailed architecture of our neural model and the strategies employed for data processing. Employing a conventional Feedforward Neural Network (FNN) structure, our model has been subjected to meticulous fine-tuning, ensuring its optimal performance. This process of fine-tuning encompasses a strategic calibration of the network's layers and hyperparameters, directed by the Optuna optimization framework. In Section 3, we will discuss the application of transfer learning techniques, emphasizing their utility in surmounting the unique data distribution challenges presented in this study. Section 4 is committed to a comprehensive analysis of our findings, including a thorough comparison with established benchmarks in the Pre-test Report. The report will culminate in Section 5, providing a complete synthesis of our research outcomes. Lastly, Section 6 will outline potential directions for future research and enhancements in this field.

II. Neural Model Architecture And Data Processing

This section will explain the design of the models' structures and the concepts of data is processed.

A. Feed-forward Neural Network Core Loss Model

Feed-forward Neural Networks (FNNs) are among the most basic and commonly employed types of artificial neural networks, demonstrating their effectiveness in addressing complex problems involving multiple variables and nonlinear regression. Figure 1 shows the structure of our model. The input layer takes in three types of parameters: frequency 'f', magnetic flux density 'B_0' to 'B_{1023}' for each point in a cycle, and

temperature '*T*'. The output layer produces one parameter: the loss per unit volume 'P_{CV}'. In between, the data passes through 'n' hidden layers, each with a different number of neurons. The exact numbers of neurons in these layers are determined using the Optuna[999] tool, which will be explained in more detail in the following sections.

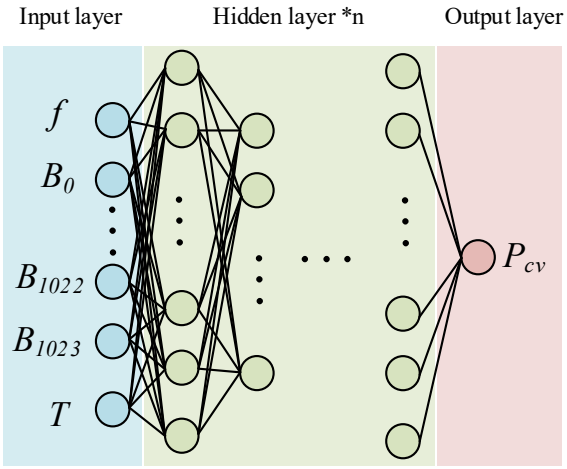

Fig. 1 A schematic diagram of the FNN architecture.

This model is synthesized and trained using PyTorch. It employs ReLU as the activation function to learn non-linearity and uses mean squared error (MSE) of logarithmic core loss as the loss function to ensure uniform performance across the entire operational range. The training optimizer selected is Adam, with an exponentially decaying learning rate strategy implemented for better model convergence. Additionally, the data is split into 60% for training, 20% for validation, and 20% for testing to assess the model's performance.

B. Optuna

Hyperparameter tuning is one of the most tedious tasks in machine learning projects. With the widespread adoption of deep learning methods, its complexity continues to increase, and the need for an effective automated hyperparameter tuning framework is more urgent than ever. To meet the demand for finding better hyperparameter configurations, our team has utilized the Optuna tool.

Optuna is an open-source Python library designed for hyperparameter optimization. Its principles are based on techniques such as Bayesian Optimization and Tree-structured Parzen Estimator (TPE) [8]. Bayesian Optimization involves constructing a probability model for the objective function and then selecting points in the search space that are most likely to improve performance. TPE is a variant of Bayesian Optimization that efficiently explores and exploits the search space using a tree-like structure.

Optuna uses Bayesian Optimization to select the optimal hyperparameter combination by defining the objective function, search space, and evaluation criteria. In addition to this, Optuna provides various built-in optimization algorithms and visualization tools, allowing us to easily conduct experiments and analysis for hyperparameter optimization. This capability has enabled us to achieve optimal performance in the project focused on predicting iron core losses.

Table.1 presents the custom-defined model hidden layers and neurons, and the configurations identified using Optuna. Table.2 compares the average percentage errors in predicting the iron core losses for ten different materials between these two configurations.

Table.1 A comparison chart of different model settings.

Model setting	Customize	Optuna
Number of hidden layers	3	5
Number of neurons in each hidden layer	15,15,15	65,55,116,40,123

Table.2 Prediction of average percentage error in core losses for ten types of materials using different models (%).

Material	3C90	3C94	3E6	3F4	77
Customize	7.28	6.1	9.48	10.31	7.16
Optuna	4.33	5.19	7.21	8.52	4.17
Material	78	N27	N30	N49	N87
Customize	7.36	9.23	5.62	7.27	11.51
Optuna	3.31	4	3.52	6.63	7.46

After obtaining the results of using Optuna to search for the number of hidden layers and neuron counts, we observed significant differences in the average percentage errors for iron core losses among ten different materials between the custom-defined and Optuna-searched hyperparameter configurations. Consequently, we further attempted to explore the initial value of the learning rate, the decay period of the learning rate, and the decay ratio of the learning rate. Table.3 presents our original hyperparameter settings alongside those identified using Optuna, with the model layers and neuron counts adopting the configuration found in Table 1 through Optuna. Table.4 compares the configurations of these two settings in predicting the average percentage errors for iron core losses of the final five materials.

Table.3 Hyperparameter settings

Hyperparameter	Customize	Optuna
Epoch	2000	2000
Decay_epoch	300	423
Decay_ratio	0.5	0.458
Initial learning rate	0.02	0.005

Table.4 Prediction of average percentage error in core losses for five types of materials in final using different hyperparameters (%)

Average percentage error	Material A	Material B	Material C	Material D	Material E	AVG
Customize	14.35	3.29	5.94	12.11	10.05	9.15
Optuna	5.57	2.17	2.78	11.61	3.59	5.14

C. Optimizing Data with Z-Score Normalization

Input data is processed using the Z-Score method (as defined in Equation 2, where 'x' represents an individual input value, 'μ' is the mean of the entire data set, 'σ' is the standard deviation of the entire data set, and 'Z' is the standardized value). The purpose of this method is to scale input variables to a standard range. This process helps in reducing the scale differences between different features, making algorithms more effective and quicker to converge. It improves the model's accuracy and performance, particularly in datasets with variables of varying scales and units, ensuring that each feature contributes equally to the analysis.

$$Z = \frac{x - \mu}{\sigma} \tag{2}$$

III. Transfer Learning

This section will explain why use Transfer Learning and provide an introduction to its concepts.

A. Data analysis

Table 5 is the dataset for the final evaluation. .There's a noticeable difference in the amount of training and testing data for Materials A and D. Material A has 2432 data points for training compared to 7651 for testing. Material D has a mere 580 points for training while there are 7299 for testing. This large discrepancy could lead to the model not having enough data to learn important features during training, which may negatively impact its performance on data it hasn't seen before.

TABLE 5 The dataset for the final evaluation

Number of Data Points	Material A	Material B	Material C	Material D	Material E
Training Data	2432	7400	5357	580	2013
Testing Data	7651	3172	5357	7299	3738

However, transfer learning allows the model to apply knowledge gained from similar tasks to better adapt to the test set. Pre-trained models are beneficial because they have already learned valuable features from a large dataset, providing a strong starting point even when there is a smaller amount of training data. The key principles of transfer learning methods are illustrated in Fig. 2.

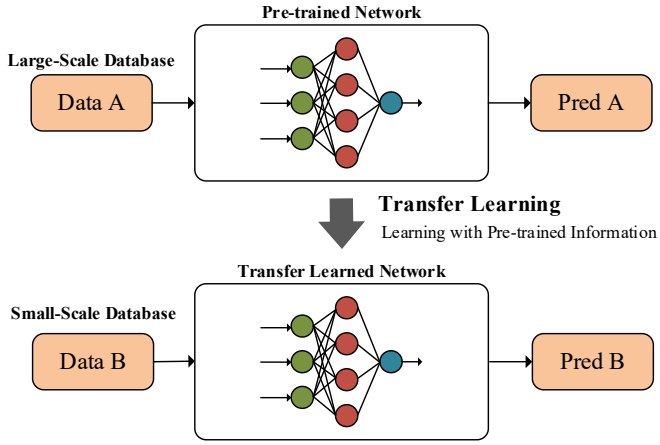

Fig. 2 The process of transfer learning: a network pre-trained on a large database is adapted to make predictions on a smaller database.

B. Decide Pre-train Models

When faced with a limited amount of training data, the filtering of pre-training data becomes crucial. In such scenarios, the model may be at risk of overfitting, relying too heavily on a limited dataset and struggling to generalize to new, similar conditions. To overcome this challenge, careful selection of pre-training data that effectively represents the target application scenarios is essential.

This report presents a method for pre-training data selection by identifying B H curves that closely resemble those of the target training model under specific conditions of temperature, frequency, and loss. The recognition method employs the Dynamic Time Warping (DTW) algorithm [9], which effectively calculates the similarity between two B H curves. As shown in Fig 3, lower DTW values indicate higher waveform similarity.

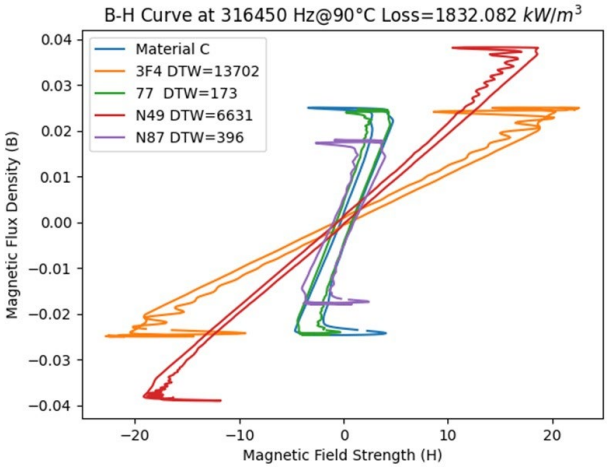

Fig. 3 Under specific conditions, the similarity of the BH curve between Material C and other materials.

According to Table 6, we can clearly observe the mutual relationship between the pre-trained model and the training model. These values are obtained through the DTW algorithm to determine which material's B-H curve is most similar under

specific conditions. Through statistical analysis, we obtain the data in the table. A higher value indicates a higher degree of similarity between the pre-training dataset and the training dataset. This metric serves as a basis for filtering the data used for training.

Table.6 The number of similar BH curve among different materials.

Material	3C90	3C94	3E6	3F4	77
A	75	132	294	1	341
B	93	223	1934	0	792
C	263	386	152	12	728
D	47	11	0	28	3
E	11	1	0	152	5
Material	78	N27	N30	N49	N87
A	256	181	348	5	98
B	528	259	2241	5	217
C	808	620	446	60	444
D	21	37	9	216	26
E	17	30	2	462	6

Ultimately, we opted for the materials with the highest metrics as our pre-training dataset. Thus, Materials A and B will use N30 as the pre-training model, Material C will use 78, and Materials D and E will use N49 as the training model. The architecture of the pre-trained models is as described in Section 2, resulting in three different pre-trained models to be used for training the five materials tested in this competition.

C. Pre-train Result

In this section, we assessed the effectiveness of pre-training across different materials. The findings, detailed in Table 7, indicated that transfer learning significantly benefited Materials C, D, and E, reducing prediction errors. However, for Materials A and B, pre-training did not yield positive results. Material A showed minimal improvement, likely due to its unique characteristics compared to other materials in the dataset. For Material B, which had ample training data, pre-training adversely affected its learning process. Based on these observations, we strategically applied transfer learning only to Materials C, D, and E, while avoiding it for Materials A and B. This decision highlights the need for a tailored approach in applying advanced techniques like transfer learning, depending on each material's specific traits and data availability.

TABLE 7 Comparison of Core Loss Prediction Error with/without Pre-training (%)

Material	Material A	Material B	Material C	Material D	Material E
Without Pretrain	4.635	1.94	3.71	11	4.35
With Pretrain	5.29	2.76	2.65	9.33	3.62

IV. RESULT

In this section, we present a comprehensive analysis of our core loss prediction model's performance. Our evaluation is centered on assessing the accuracy and reliability of the model across different materials, ranging from Material A to Material E. The results are quantified in terms of average error rates, percentiles of relative cumulative frequency, and maximum errors. Additionally, we provide insights into the uniformity of the model's architecture by examining the number of parameters for each material model. This detailed analysis aims to offer a clear understanding of the model's efficacy in predicting core losses under varied conditions.

A. Result

Table 8 presents the error data for Materials A to E, featuring metrics such as the average error, the 95th percentile of relative cumulative frequency (95-PRCT), the 99th percentile (99-PRCT), and the maximum error (Max). Except for Material D, the average errors for the other materials do not exceed 5%. Table 9 lists the Number of Parameters for each material model, which is uniform at 86,728 across all models due to the identical model architecture employed.

TABLE 8 Core Loss Prediction Error

Material	Material A	Material B	Material C	Material D	Material E
Avg	4.635	1.94	2.65	9.33	3.62
95-PRCT	14	4.84	6.95	30.03	12.26
99-PRCT	29.58	12.156	12.69	41.77	22.97
Max	59.56	58.049	49.05	46.6	49.56

TABLE 9 Number of Parameters for Each Material

Material	*A*	*B*	*C*	*D*	*E*
Number of Parameters	*86728*	*86728*	*86728*	*86728*	*86728*

B. Comparison of Errors Relative to the Pre-test Report

Table 10 presents the progress from midterm to final evaluations across ten materials, showing significant improvements in the final scores. The initial approach during the midterm, which utilized LSTM to forecast the magnetic field intensity H and subsequently calculated losses via the BH curve area, did not yield the expected results. This shortfall might be due to the limited complexity of the LSTM, leading to less precise predictions of H, and the inherent inaccuracies when correlating the BH curve area with actual volumetric core losses. Ultimately, transitioning to a conventional FNN model, although more parameter-intensive, facilitated a more accurate direct prediction of losses than the indirect BH curve area method.

Table.10 Core Loss Prediction Error: Midterm vs. Final Comparison (%).

Material	3C90	3C94	3E6	3F4	77
Pre-test Report	48.58	46.61	23.99	112.10	49.45
Final-test Report	12.2	12.81	16.55	21.2	11.81
Material	78	N27	N30	N49	N87
Pre-test Report	49.45	41.13	19.58	173.50	32.91
Final-test Report	9.31	10.18	8.34	17.23	20.2

The results outlined in this report underscore the effectiveness of our deep neural network model in predicting core losses across a diverse range of materials. Notably, except for Material D, the model maintains an average error rate below 5%, signifying its high accuracy. The uniformity in the number of parameters across all material models, standing at 86,728, reflects a consistent approach in the model's architecture. This consistency further reinforces the reliability of our findings. The data presented in Tables 8 and 9 collectively demonstrate the model's robustness and precision, making it a valuable tool in the field of core loss prediction.

V. Conclusion

In conclusion, this report successfully illustrates the effective use of a Feedforward Neural Network (FNN) in predicting core loss. Through the meticulous fine-tuning of the network architecture using Optuna, a robust model comprising five hidden layers with varying neuron counts (65, 55, 116, 40, and 123) was developed. The calibration of hyperparameters, including a decay epoch of 423, a decay ratio of 0.458, and an initial learning rate of 0.005, played a significant role in enhancing the model's performance. The application of Z-score normalization was particularly effective in improving accuracy and efficiency.

Addressing the challenge of limited training data, the incorporation of transfer learning and the Dynamic Time Warping (DTW) algorithm proved instrumental. This approach enabled the identification of the most relevant pre-test materials for pre-training, optimizing the model's learning process. However, it's noteworthy that pre-training did not improve outcomes for Materials A and B, indicating that the effectiveness of transfer learning can vary depending on the material type. Consequently, transfer learning was selectively applied to Materials C, D, and E.

The data partitioning strategy, with 60% allocated for training, 20% for validation, and 20% for testing, was effective in achieving an average error below 5% for most materials. The notable exception was Material D, which demonstrated the model's limitations and potential areas for future improvement. Additionally, Material B exhibited the highest error rate at 58.049%, indicating a need for further refinement in the model's approach to specific material types.

VI . Future work

Future research will focus on enhancing accuracy for challenging materials, like the high error rates seen in Material B (58.049%). This involves integrating diverse datasets with a wider range of material characteristics and developing specialized algorithms for materials like Material D.

Given the variable success of transfer learning, future studies will explore sophisticated methods to identify material similarities and extend transfer learning to more materials. Improving pre-training techniques, particularly for less responsive materials, is also a key goal, possibly through advanced feature extraction methods.

To address limited training data, the plan includes expanding the dataset through synthetic data generation or additional real-world data collection. Parallel to this, there will be a continuous effort in optimizing hyperparameters like learning rates and neuron configurations, enhancing model performance and accuracy.

Future models will benefit from integrating multi-modal data sources, such as thermal and electrical properties. Developing ensemble models that combine the FNN with other neural network types may yield better results for challenging materials.

These strategies aim to overcome current limitations, significantly improving the model's accuracy and applicability in core loss prediction, and enhancing its utility in industrial applications.

References

[1] 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," in IEEE Transactions on Industry Applications, vol. 37, no. 4, pp. 1055-1061, July-Aug. 2001, doi: 10.1109/28.936396

[2] J. Li, T. Abdallah, and C. R. Sullivan, “Improved Calculation of Core Loss With Nonsinusoidal Waveforms”, IEEE Industry Applications Society Annual Meeting, Oct. 2001, pp. 2203–2210.

[3] K. Venkatachalani, 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 Electronics, 2002, pg. 36-41. Reference 181.

[4] J. Muhlethaler, J. Biela, J.W. Kolar, A. Ecklebe, "Improved Core-Loss Calculation for Magnetic Components Employed in Power Electronic Systems," Power Electronics, IEEE Transactions on , vol.27, no.2, pp.964-973, Feb. 2012.

[5] D. Lin, P. Zhou, W. N. Fu, Z. Badics and Z. J. Cendes, "A dynamic core loss model for soft ferromagnetic and power ferrite materials in transient finite element analysis," in IEEE Transactions on Magnetics, vol. 40, no. 2, pp. 1318-1321, March 2004, doi: 10.1109/TMAG.2004.825025.

[6] J. Mühlethaler, J. Biela, J. W Kolar, and A. Ecklebe, "Core Losses Under the DC Bias Condition Based on Steinmetz Parameters," in IEEE Transactions on Power Electronics, vol. 27, no. 2, pp. 953-963, FEBRUARY 2012, doi: /TPEL.2011.2160971.

[7] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural Networks, vol. 61, pp. 85–117, jan 2015.

[8] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: a Next-Generation Hyperparameter Optimization Framework,” Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2019.

[9] M. Müller, "Dynamic time warping," in *Information retrieval for music and motion*, pp. 69-84, 2007.