MagNet Challenge 2023 - University of Bristol

Lizhong Zhang, Tom McKeague, Binyu Cui, Navid Rasekh,
Jun Wang (Academic Advisor), Song Liu (Academic Advisor), Alfonso Martinez (Industry Advisor)
{ts20212, rd20871, bc16461, ck19302, Jun.Wang, Song.Liu}@bristol.ac.uk; alfonso.martinez@we-online.com

***Abstract*—This report describes the work done by the University of Bristol Team in response to the MagNet Challenge 2023.**

Keywords—MagNet Challenge, machine learning, neural networks

I. Introduction

Nowadays, magnetic components are involved in most power electronics applications for functionality and filtering purposes. It's typically known to be the least efficient component that has a massive impact on system performance and efficiency regarding the size, weight and power loss factor [1][2]. However, an accurate core loss model for magnetic components that is based on the first principle remains elusive due to the non-linear feature of the magnetic material and other intercoupled factors such as DC-bias condition. Numerous research studies have been carried out to factor in the external parameters contributing to magnetic loss under nonsinusoidal excitations. The Steinmetz equation (SE), shown in (1), is widely accepted as an empirical model to calculate core loss under sinusoidal excitation. The k,α and β are the SE parameters which can be calculated by substituting the measured core loss value with the corresponding frequency and peak flux density values. This results in diminished precision of the equation due to the SE parameters demonstrating inconsistent performance across different frequency ranges.

$$P_{loss} = kf^{\alpha}B^{\beta} \tag{1}$$

To enhance the universality and accuracy of the SE, the improved generalised Steinmetz equation (iGSE), as a modified solution, has been proposed based on SE for calculating core loss for arbitrary flux waveform under zero DC-biased condition [3]. The core concept of iGSE is to divide the complex waveform into individual B-H loops and calculate the loss respectively. However, these models generally face limitations in accuracy, particularly with certain waveform types, and tend to overlook the effects of temperature. Another approach called 'Loss map' is proposed in [4][5] for incorporating the pre-magnetization effects. To begin with, the operating state of one magnetic component can be described by three variables, the pre-magnetization state, the magnetic flux density swing and the flux density change rate. By measuring the B-H loop at various preset operating points, a core loss profile can be produced to cover all the operating conditions and used directly as a look-up table in further in-situ experiments.

In recent years, methods using neural networks and other machine learning techniques have demonstrated excellent results in addressing nonlinear regression problems and forecasting time series data (e.g., image & voice identification) [6]. This technique could also be applied in core loss predicting and has proven to be more precise than the classic modelling based on the Steinmetz equation [7][8].

Three neural network models have been proposed and discussed in [7]. There is the 'scalar-to-scalar' model such as FNN which acts similarly to the Steinmetz equation that uses parameters such as flux density and frequency to directly predict the power loss. Unlike the limitations of SE, the FNN-based model is proven to have a higher accuracy rate across the frequency range while covering external influencing factors such as temperature. One drawback of this model is that different models have to be trained according to the excitation waveform type which the scalar could not fully represent. On the other hand, instead of using magnetic scalars, the 'sequence-to-sequence' model such as the transformer model [9] introduces the complete excitation waveform to the training process and predicts the magnetic response. While solving the issue of producing corresponding models to waveform types, the 'sequence-to-sequence' model normally introduces an enormous amount of parameters in both the input and output sides of the model which would lead to a longer training process and higher requirement for the training platform.

To leverage the strengths of the aforementioned models, the Long Short-Term Memory (LSTM) network which is a 'sequence-to-scalars' model has been chosen as our model for participation in the MagNet challenge. The LSTM network excels in processing regression problems with sequential input due to its key feature of capturing long-term dependencies in data and overcoming the short-memory issue prevalent in standard RNN models [10]. These characteristics are ideally suited to fulfil the requirements for processing and analysing the time series data relationships, such as those between B(t) and H(t).

In response to the MagNet Challenge 2023, this report details the approaches applied by the University of Bristol team for the competition. Section II gives an overview of the challenge. Section III introduces the machine learning framework and detailed considerations. Section IV shows the results and a summary of our models.

II. The Goal and Challenges

The aim of this project is to yield a "prediction model" for one magnetic core material, which takes in three inputs, B, f and T and outputs one volumetric loss density value. To achieve this aim, a machine learning process is expected to learn from the large database provided, which is experimentally measured loss and waveforms. This database is treated as the "ground truth" in

this work. Hence the top-level idea of this work is illustrated in Fig. 1

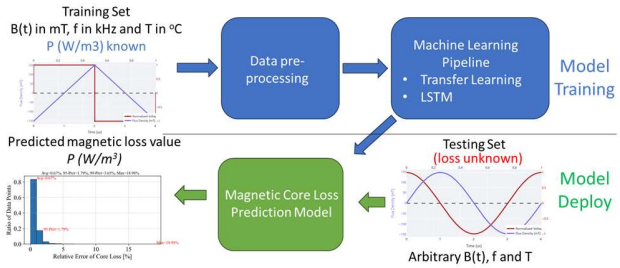

Fig. 1. The goal

There are a few challenges embedded in the training and testing datasets provided to us, which reflecs real-life user senarios. As illustrated in Fig. 2, there is a large-scale database of ten magnetic materials provided as the solid ground and starting point of this work. This large-scale database can serve as the foundation for the pre-training in a transfer learning process. Reflecting the real-world senario to adapt a trained base model to a new material, the data available for five unknown materials are given with a small and unbalanced training set. The final testing data is given with inputs only to test out the model predictions.

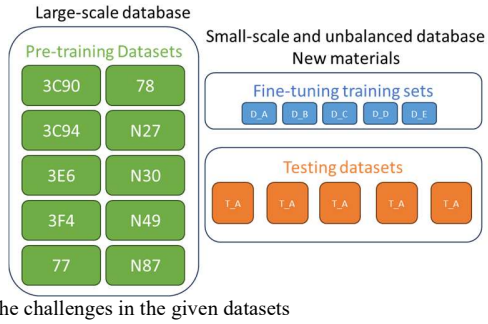

Fig. 2. The challenges in the given datasets

III. Machine Learning Framework

A. *Overall structure/pipeline with Transfer Learning*

Given the rich training set for the ten known materials and limited data for the five unknown materials, this work has adapted a transfer learning pipeline, which contains a pre-training stage and a fine-tuning stage. This pipeline is illustrated in Fig. 3. Starting from zero in the pre-training stage, the machine learning model is trained by a large dataset of one particular material. As the most commonly-used and perhaps the most representative material in engineering, the training data of material 3C90 is used to generate a "fundamental model" as a generalised model to extract the key physical patterns for the general task of predicting the core loss value given the input data. The fundamental model is then tuned in the second-stage of pre-training against the ten materials' training set, which yields 10 "base models" $\{M_i, i = 1…10\}$.

In the fine-tuning stage, one of the base model is selected and tuned specifically for one particular unknown material (e.g. Material A) in the final test data. This stage starts with a selection logic to pick out the best tuned base models based on the minimal average error that one base model yields on the testing data, which can be considered as identifying one material out of the original ten materials that is most similar to Material A. At the end, a fine-tuned model for Material A is generated for deployment.

This pipeline is designed to make the most out of the rich training set of the original ten materials and transfer the learned knowledge into the fine-tuning step, which is a solution to the limited dataset for the testing data/materials. This idea is inspired from a machine learning concept called "meta learning". The "meta" parameters of a neural network are pretrained using a similar material (3C90) and adapted to the target materials with further finetuning.

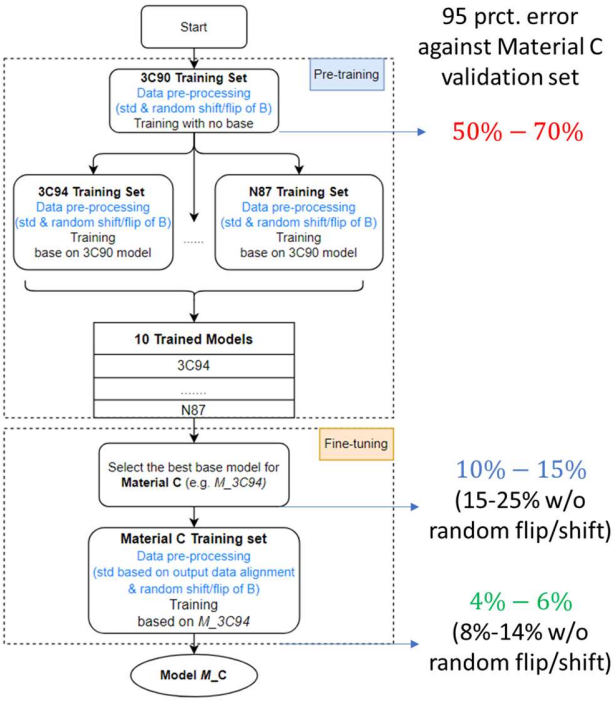

Fig. 3. Overall pipeline with transfer learning (example with 3C90 as the fundamental base and Material C as the target)

For the training process, the data are split into training set, validation set and a test set on a 70-20-10 basis. The split is performed randomly to minimize bias.

B. *Data pre-processing*

As marked in blue texts in Fig. 3, the training data has gone through a data pre-processing process before each training task takes place.

Data standardization

The input variables temperature and frequency are linearly rescaled into values in the range of [0, 1]. The *B* waveform and the loss density are also standardized so values are in the range of [-1, 1]. This procedure is to retain the polarity information and the linear correlation between the primary input (B waveform) and the single model output (loss density) for better LSTM waveform feature extraction. We observe that this data standardization process can enhance the model performance, avoids numerical instabilities and improves model performance. For the LSTM, data standardization can improve the convergence in the gradient descent optimization process. In the case of standardising the B waveform, it can be considered as extracting the shape of the waveform and minimize the numerical impacts of the magnitude of data. The standardization process is illustrated in Fig. 4, where the raw data is processed

into standardized data and stored together with the linear standardization coefficients (k, b) for each case. The scaling/standardization coefficients are determined for each material based the range of data, e.g. max value.

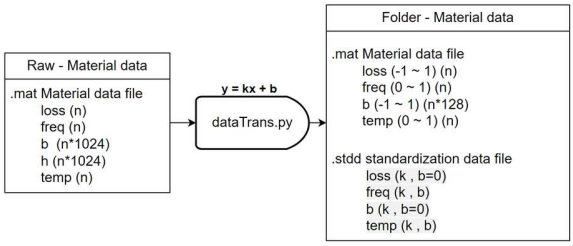

Fig. 4. Data standardization process (linear rescaling through $y=kx+b$)

Random shift/flip of *B* waveform

To enhance the generalisation and avoid overfitting to the time-sequence waveform data, a "random shift/flip" operation is implemented on the B waveforms before they enter the training stage. The assumption is that shifting horizontally (i.e. applying a phase shift) and/or flipping (vertically and horizontally) the B waveform will not change the core loss density associated with this waveform – based on domain knowledge, this assumption should be valid given the core loss density is an averaged value over one whole cycle of a periodic B waveform repeating itself once every 1/f second. This technique is similar to duplicating training images by rotating them in image recognition tasks. By injecting artificially rotated images, the neural network is trained to capture rotation invariant information from images. However, in our case, a rotation (pivoting) of the B waveform, instead of shifting or fliping, will invalid the euqal loss assumption. Note the phase shifting is similar to the approach in [11], while the flipping is an original contribution of this work.

To implement this operation, a function is applied on B waveforms to shift and/or flip them randomly before feeding them into the LSTM process as the sequential data input. Based on observations, this technique can significantly improve the model accuracy and reduce overfitting to the sequence of the signal.

Downsampling

The original 1024-point B waveform is downsampled to 128-point data based on a trial-and-error approach to reduce the computation need in the training process and model size while it still maintains a level of accuracy without missing key patterns/details. This operation may also reduce the overfitting on irrelevant/insignificant details contained in the waveforms.

Output data alignment for data standardization in the fine-tuning stage

To cope with the small-training-set problem, a special "output data alignment" technique is applied in the fine-tuning stage as a data standardisation approach. Given the objective in the fine tuning stage is to minimise the modification of the base model and unwanted overfitting towards the small training set, the output data (core loss density) to be fed into the training process is standardized in a different way – they are rescaled in a manner to align their extracted features with those of the original training data, to enhance the fine-tuned model. In order to correct the output shift, we apply shift and rescaling to y to match the distributions in the source/training set.

C. LSTM Neural Network Architecture

An LSTM + feedforward neural network (FNN) framework is applied as the fundamental machine learning approach. As shown in Fig. 5, LSTM is applied to deal with the ***B*** waveform only, excluding the temperature and frequency. This design enhances the feature extraction and pattern recognition of the waveforms to be better transferred to other models in the transfer learning architecture. The neural network generates one scalar, the volumetric magnetic loss, as the output – the whole process is a sequence-to-scalar model.

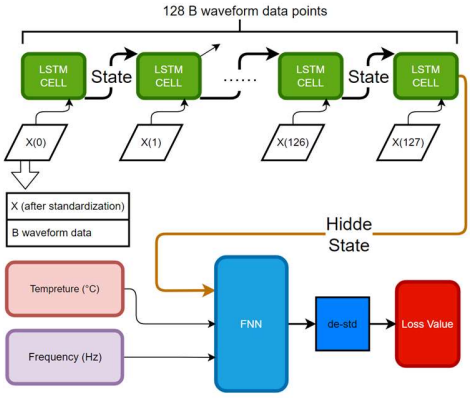

Fig. 5. Example of LSTM neural network structure

The loss function for the optimization process of the neural network is defined to minimize the relative error.

D. Data post-processing

Data de-standardization

Following the data standardization, we apply a reverse process to translated back the loss density value to the original scale at the end of the prediction pipeline.

Averaging output with random shift

When we deploy the model, shifting/flipping the ***B*** waveform can still lead to variations of the predicted loss. To take this factor into account, for each entry of the test data, the B waveform is processed into 100 (adjustable) instances with different phase shift/flips, which yields 100 predicted volumetric loss values. These values are averaged in a post-processing step to obtain the final prediction for this entry.

E. Data imputation

To cope with the cases with limited and unbalanced training set, we applied a data imputation approach to artificially expand the training set in these cases. For example, the training set of Material D only contains 580 entries, with an unbalanced split between sinusoidal, trangular and trapezoidal waveforms (e.g. 145/400/35). However, the testing data of Material D is 7299 entries with a 61/2247/4991 split – the testing data has a large portion with trapezoidal waves, while the training set has very limited data for this case. To compensate this mismatch, the trapezoidal data in the training set is artificially duplicated and expanded to match the waveform split in the test set as much as possible, which leads to a 145/2000/700 split.

IV. Results and Summary

Based on the above machine learning framework, five models are produced for the five unknown materials. The performance of these models are evaluated against the validation

sets in the five cases, with the results shown in Fig. 6. Although the average of 95-percentile-error for the 10 known materials can achieve around 2.5%, the model's performance deteriorated in the cases of the 5 unknown materials due to the limited training sets. Material D is particularly problematic due to only 580 entries of training data available, which leads to a 22.04% 95-percentile-error. The average 95-percentile-error of the other four unknown materials achieves a 6.6%.

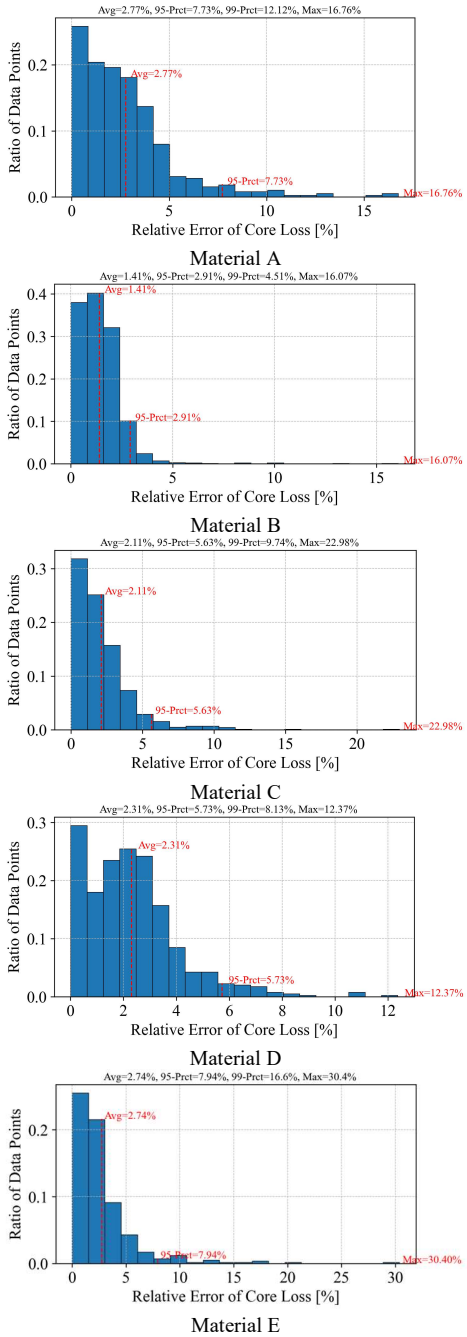

Fig. 6. Error distribution against the validation set for the target five unknown materials

The size of the five models are listed below in terms of number of parameters and file sizes. They all have the same size since our model architecture is fixed when training the base mode.

Table I Model Size

	Material A	Material B	Material C	Material D	Material E
Number of parameters	15,653	90,653	90,653	16,449	16,449
Model Size	361 KB	361 KB	361 KB	70 KB	70 KB

Overall, this work has developed and demonstrated a machine learning pipeline in response to the MagNet Challenge 2023. An LSTM + FNN structure is applied as the core machine learning approach. Several techniques are applied in the data pro-processing stage to enhance the model accuracy, reduce computation load and minimize bias/overfitting. For example, based on domain knowledge, a random shift/flip operation is applied on the B waveforms as pre-processing to reduce the sensitivity on the sequence of the waveform data being fed into the LSTM step for feature extraction. A special data standardization method to align the output data is applied in the fine-tuning stage to minimize the modification to the base model while cope with the small training set.

Excellent accuracy has been achieved for the 10 known materials with large-scale datasets. To tackle the challenges caused by the small datasets for the five unknown materials, special standardization and data imputation techniques are applied.

Further improvements are being tested to further balance the model performance and size, such as data imputation and model distilling methods.

REFERENCES

[1] K. Venkatachalani, C. R. Sullivan, T. Abdallah, and H. Tacca, "Accurate prediction of ferrite core loss with non-sinusoidal waveforms using only Steinmetz parameters," in Proc. IEEE Workshop Comput. Power Electron., 2002, pp. 36–41.

[2] C. R. Sullivan, "High frequency core and winding loss modeling," in Proc. IEEE Int. Electric Mach. Drives Conf., 2013, pp. 1482–1499.

[3] J. Li, T. Abdallah, and C. Sullivan, "Improved calculation of core loss with nonsinusoidal waveforms," in Proc. Conf. Rec. IEEE Ind. Appl. Conf. 36th IAS Annu. Meeting, vol. 4, 2001, pp. 2203–2210.

[4] J. Wang, K. J. Dagan, X. Yuan, W. Wang and P. H. Mellor, "A practical approach for core loss estimation of a high-current gapped in-ductor in PWM converters with a user-friendly loss map," *in IEEE Transactions on Power Electronics*, vol. 34, no. 6, pp. 5697-5710, June 2019.

[5] S. Iyasu, T. Shimizu and K. Ishii, "A novel iron loss calculation method on power converters based on dynamic minor loop," *2005 European Conference on Power Electronics and Applications*, Dresden, Germany, 2005, pp. 9 pp.-P.10.

[6] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Netw., vol. 61, pp. 85–117, Jan. 2015.

[7] H. Li et al., "How MagNet: Machine Learning Framework for Modeling Power Magnetic Material Characteristics," in *IEEE Transactions on Power Electronics*, vol. 38, no. 12, pp. 15829-15853, Dec. 2023, doi: 10.1109/TPEL.2023.3309232.

[8] N. Rasekh, J. Wang and X. Yuan, "Artificial Neural Network Aided Loss Maps for Inductors and Transformers," in *IEEE Open Journal of Power Electronics*, vol. 3, pp. 886-898, 2022, doi: 10.1109/OJPEL.2022.3223936.

[9] A. Vaswani *et al.*, "Attention Is All You Need," *arXiv.org*, Dec. 05, 2017. https://arxiv.org/abs/1706.03762

[10] X. Li and X. Wu, "Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2015, pp. 4520–4524.

[11] H. Li et al., "How MagNet: Machine Learning Framework for Modeling Power Magnetic Material Characteristics," in IEEE Transactions on Power Electronics, 2023, pp. 15829-15853.