A Numerical Twin Model for the Coupled Field Analysis of TEAM Workshop Problem 36

P. Di Barba[1], M. E. Mognaschi[1], A. M. Cavazzini[2], M. Ciofani[2], F. Dughiero[2], M. Forzan[2], M. Lazzarin[2], A. Marconi[2], D. A. Lowther[3], *Life Fellow, IEEE*, and J. K. Sykulski[4], *Life Fellow, IEEE*

[1]Department of Electrical, Computer and Biomedical Engineering, University of Pavia, 27100 Pavia, Italy
[2]Department of Industrial Engineering, University of Padua, 35131 Padua, Italy
[3]Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada
[4]School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K.

A numerical twin model for the magneto-thermal analysis of an induction heating device is proposed. The non-linearity of magnetic permeability against temperature—which characterizes the workpiece—is captured by the model, while the use of a convolutional neural network (CNN), trained by a number of finite-element (FE) analyses, makes it possible to solve the following inverse problem: given a temperature map in the workpiece section, identify current and relevant frequency in the inductor coil, as well as the time instant at which the map refers to. The testing electromagnetic analysis method (TEAM) problem 36 is considered as the case study.

***Index Terms*— Convolutional neural network (CNN), coupled problems, testing electromagnetic analysis method (TEAM) problem.**

I. Introduction

In THIS article, a study is presented dealing with the feasibility of solving direct and inverse coupled field problems by resorting to deep learning techniques and surrogate models. Nowadays, surrogate models based on deep neural networks are increasingly used in computational electromagnetics (EMs); to give some examples, deep learning techniques have been used for field analysis in single-physics [1], [2] and multiphysics [3] domains, while fully unsupervised approaches have been shown to be effective in solving problems of material property identification [4]. The coupled field problem deals with the mass heating of a steel billet that is a "classical" induction heating application, where the calculation of the time-dependent temperature field results from a coupled magneto-thermal analysis, traditionally carried out by numerical methods such as finite elements (FEs) or finite differences. Recently, the objective of this investigation has been added to the testing electromagnetic analysis method (TEAM) benchmark problems as "Problem 36: Multiphysics Field Analysis of an Induction Heating Device" [5], [6].

The authors have already assessed, for the same benchmark problem, the possibility of applying convolutional neural networks (CNNs), both on a pretrained neural network—notably GoogLeNet—and on a neural network trained from scratch [7]. In the previous work, CNNs were trained by a limited number of FE solutions, with a database that incorporates two data sets: a reduced one with 120 FE solutions and a larger one with 1654 solutions, both calculated through a simplified approach to consider the temperature dependence and non-linearities of material properties. In particular, the simplified approach was based on the assumption of a two-level linearized model: taking the magnetic and electrical properties of the billet constant values below and above the Curie temperature, respectively [7].

The training data were the input variables, current density, and frequency of the inductor supply and the resulting duration of the heating process to reach 1200 °C at least at one point on the billet and the corresponding final temperature map sampled on a regular grid of 32 × 12 points [7]. In contrast, in this work, training data have been obtained from FE coupled problems developed using a full-transient thermal analysis with non-linear material properties: the database of results incorporates the temperature maps, sampled on a 30 × 60 regularly spaced grid, as calculated during the time transient; in particular, 24 temperature maps, relevant to 24 time instants, have been stored for each FE solution. The complete training set contains about 40 000 temperature maps that arise from 1600 combinations of input data, frequency, and current, thus preserving the knowledge of the transient evolution of temperature for each case.

II. TEAM 36 Multiphysics Problem: The Field Model

The use of numerical models for the design of induction heating systems has been characterized by a rapid evolution in recent years; in fact, they have become a dominant method of designing electrothermal systems [8]. Moreover, numerical modeling has made it possible to significantly reduce design times and costs, as before the introduction of advanced and reliable software, the design depended on the expertise of the designer who could rely only on formulae based on simplified analytical models. In the TEAM 36 problem, an induction heating problem related to coupled EM and

Manuscript received 4 November 2022; revised 10 January 2023; accepted 14 January 2023. Date of publication 23 January 2023; date of current version 25 April 2023. Corresponding author: M. E. Mognaschi (e-mail: eve.mognaschi@unipv.it).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMAG.2023.3238767.

Digital Object Identifier 10.1109/TMAG.2023.3238767

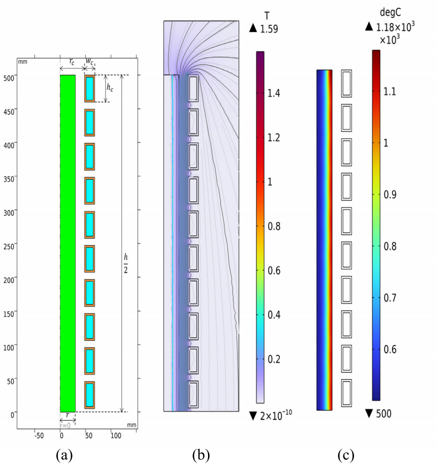

Fig. 1. (a) Geometry of the system $h/2 = 50$ cm, $r = 3$ cm, $h_c = 2$ cm, and $w_c = 1$ cm; (b) magnetic field map; and (c) thermal map at $t = 7$ s, $I = 6$ kA, and $f = 6$ kHz.

thermal calculations is proposed [5] and is used in this article to evaluate the possibility of using metamodels based on neural networks for coupled problem. The device under study is composed of a solenoidal inductor coil and a cylindrical steel billet, coaxially located with respect to the inductor and of the same axial length, presented in Fig. 1 with reference to a cylindrical coordinate system $\{\rho, \phi, z\}$.

The billet is made of the C45 steel, whose material properties, specific heat capacity, thermal conductivity, electrical conductivity, mass density, and permeability, as functions of temperature and/or magnetic field intensity, are given in [6].

An FE model of the device was derived using Comsol Multiphysics [9] and, due to the symmetries, it was possible to apply a 2-D axisymmetric model to half of the inductor and steel billet as well as the air for the coupled solution. The following boundary conditions were applied:

1) tangential flux lines along $\rho = 0$;
2) normal flux lines along $z = 0$.

In turn, the thermal domain is composed of half of the billet, with relevant boundary conditions applied to the external contour of the billet, to describe convective and radiation losses from the billet surface. The convective exchange coefficient was assumed to be equal to 7 W/m/°C and the emissivity coefficient equal to 0.8, while the room temperature was equal to 70 °C along the lateral surface of the billet ($\rho = 3$ cm) and equal to 25 °C at the end surface ($z = h/2 = 50$ cm).

The EM problem is solved in time harmonic, while the thermal one is in time-dependent conditions. The distribution of Joule's losses in the billet, as computed at each EM step, is the coupling term with the transient thermal, where it is applied as a heating source.

In turn, the resulting temperature distribution at the end of each time step is the coupling term with the EM solver, where material properties are updated accordingly to the actual temperature.

The benchmark problem refers to a specific process, where the inductor is supplied by a sinusoidal current of 3500 A (rms), at frequency $f = 2$ kHz. The heating process lasts 250 s, and the calculated temperature distributions along

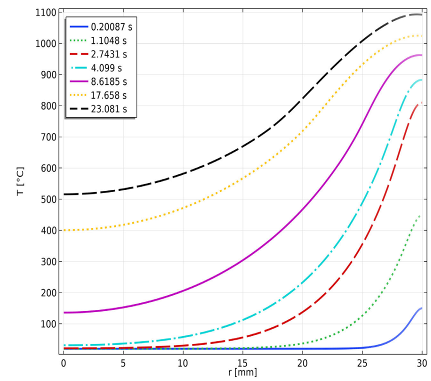

Fig. 2. Temperature distribution along the billet radius at $z = 0$, with different time instants considered.

the billet radius located on the symmetry plane (i.e., at $z = 0$) are shown in Fig. 2.

In practice, the coupled solution requires several EM solutions, which are non-linear below the Curie temperature, and thermal solutions, which also exhibit strong nonlinearities, mostly related to the specific heat and radiation losses.

Typical computational times of each coupled problem are quite high, even if calculations are performed on powerful workstations and taking advantage of parallel computing for solving the big complex and real matrix produced by the FE method. For a typical FE simulation, there are about 170 000 degrees of freedom considering both magnetic and thermal problems.

III. CNN-Based Approach to the Identification Problem

Starting from the temperature distribution in the billet cross section at a given time instant, the proposed method aims at recovering the electrical input of the inductor coil, as well as the time instant itself. Hence, an identification problem arises that reads as follows: given a matrix representing a temperature map, identify the corresponding vector of current, frequency, and time instant. More generally, the temperature map can be understood as a representative of a prescribed thermal field that should be inverted.

For solving the problem, a CNN-based approach is used. In such a case, the network is a surrogate of the inverse problem, meaning that applying an image of the temperature field as an input to the CNN, the current, frequency, and time instant of the heating process is identified without the use of an FE model.

In order to train the CNN for solving this problem, a database of field solutions, based on the FE model, was created as follows.

Along the radius r, N_r points were taken inside the billet domain, as well as N_z points on the semi-length z; by considering a given time instant t^*, a temperature matrix of dimension $[N_r \times N_z]$ was defined, whose rows corresponded to the radius coordinates, while the columns corresponded to the semi-length ones: each element was equal to the temperature value at the relevant point and at the instant t^*. This matrix can be built for each instant t_k, $k = 0, 1, \ldots, n_t$, with n_t the total number of selected instants, in order to obtain a tensor of

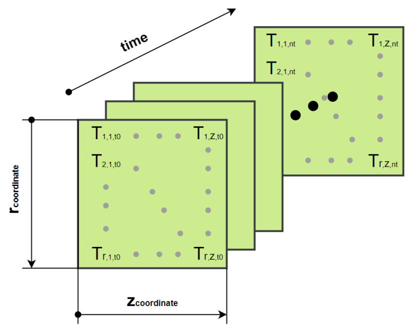

Fig. 3. Structure of the database.

dimension $[N_r \times N_z \times n_t]$ that stores the temperature values in the $N_r \times N_z$ points of the billet and for the n_t time instants, as shown in Fig. 3.

It is important to note that each analysis case, distinguished by the amplitude and the frequency of the source current flowing into the coil, exhibits a final time different to each other since the heating power is affected by the above parameters.

Furthermore, the number of intervals actually computed in each case is different because the solution time steps are automatically and adaptively set by the transient solver in Comsol (the interpolated solutions are discarded). For this reason, in order to enhance and simplify the training process of the CNN, the number of time instants stored for each case has been standardized by taking the same number of computed intervals n_t; in this way, it was possible to include only the effectively computed solutions and to build a 5-D tensor of size $[N_{\text{curr}} \times N_{\text{freq}} \times N_r \times N_z \times n_t]$, with N_{curr} and N_{freq} the numbers of the studied values of current amplitude and frequency, respectively. The last component of the proposed database is a tensor of dimension $[N_{\text{curr}} \times N_{\text{freq}} \times n_t]$, whose elements are the time instants relevant to each case, since the sampling process is carried out on the numbers of computed time steps and not on the absolute time.

The database of solutions is the training set for the CNN and comprises all the results obtained by applying a current density ranging from 2000 up to 6000 A (rms values) and frequency ranging from 2000 up to 6000 Hz. Overall 1600 transient field analyses have been performed.

The generation of the database of solutions took about 40 days of calculations distributed on three different workstations characterized by 2× Intel Xeon E5-2620 v4 processors (2.10 GHz, 16 cores, and 32 threads) and 256 GB of RAM.

For solving the identification problem, a CNN trained from scratch is used. The CNN is composed of 18 layers, in which four blocks can be highlighted. Each block is composed of a convolutional layer, a batch normalization layer, and a rectified linear unit (ReLU) function, as shown in Table I.

The ReLU function is one of the most used activation functions for CNN because it has shown good performances in training this kind of neural networks in terms of avoiding the overfitting phenomenon.

The convolutional layers are characterized by filters of size 3 × 3, while the number of filters varies from 128 to 8, depending on the block. Moreover, two average pooling layers with filter of size 2 × 2 are applied in order to obtain a more stable solution. In the final block, a dropout layer is used and a fully connected layer followed by the regression layer allows a vector of three elements to be obtained as the output.

TABLE I
CNN ARCHITECTURE

Blocks	Layers
Input	Image based input (size 30x60)
Block 1	Convolution 2D (size 3x128),
	Batch Normalization
	ReLU activation function
	Average Pooling 2D (size 2x2)
Block 2	Convolution 2D (size 3x64)
	Batch Normalization
	ReLU activation function
	Average Pooling 2D (size 2x2)
Block 3	Convolution 2D (size 3 x 32)
	Batch Normalization
	ReLU activation function
Block 4	Convolution 2D (size 3x8)
	Batch Normalization
	ReLU activation function
	Dropout (20% probability)
	Fully connected layer (3 outputs)
Output	Regression layer

The architecture of each block in Table I is typical of CNNs that have a field image as the input. The number of blocks is a tradeoff between two conflicting criteria: a low number of blocks could result in an oversimplified net despite the complexity of the field problem to solve and a high number of blocks could result in a larger dataset for training purposes.

The CNN is able to treat images as input, and here, it is used as follows: given a temperature map, the corresponding vector of current, frequency, and time instant is predicted; this mapping is a model of the inverse induction heating problem.

In terms of an error estimate, for evaluating the quality of the CNN prediction, the root-mean-square error e is calculated on N points of the validation set, namely,

$$e = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2}{\sum_{i=1}^{N} Y_i^2}} \tag{1}$$

where Y is the true value used in the finite element model (FEM) and $\hat{Y}$ is the value predicted by the CNN (both normalized in the range between -1 and 1). The error is evaluated for each of the three quantities: current, frequency, and time instant.

IV. RESULTS

The CNN is trained following roughly the 80/20 rule, i.e., 80% of the dataset for the training set and 20% for the validation set. Thus, training and validation sets have different sizes depending on the dataset, as shown in Table II. The training was performed with ADAM solver and lasted 1000 epochs.

The plane "true versus predicted values" shows the relevant level of agreement: the closer the points are to the diagonal, the more accurate the solution. The prediction for the time instant is accurate for those instants at the beginning of the process, while the final time instants are not very well predicted, as shown in Figs. 4–6 (red squares for which the true value is

TABLE II
SIZE OF TRAINING AND VALIDATION SETS

Dataset	Training set	Validation set
Dataset 1(4 time instants)	5,400	1,324
Dataset 2(9 time instants)	12,100	3,029
Dataset 3(14 time instants)	19,500	4,034
Dataset 4(19 time instants)	25,500	6,439
Dataset 5(24 time instants)	33,000	7,344

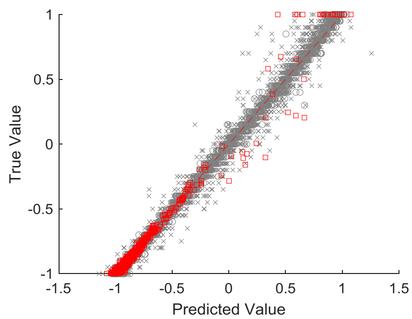

Fig. 4. True versus predicted values of current (circle), frequency (cross), and time instant (red square) for Dataset 1.

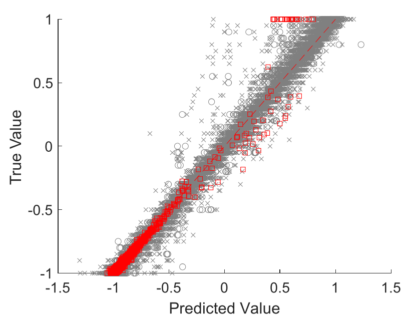

Fig. 5. True versus predicted values of current (circle), frequency (cross), and time instant (red square) for Dataset 3.

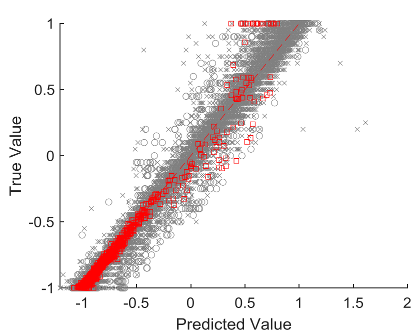

Fig. 6. True versus predicted values of current (circle), frequency (cross), and time instant (red square) for Dataset 5.

equal to 1). Hence, when the number of time instants increases, this error is lower percentagewise, considering the whole set of time instants. In Table III, an estimate of the error e, computed for the validation set according to (1) and referring to the values normalized in the range $[-1,1]$, is reported.

From Table III, it will be noted that the current prediction is slightly more accurate than the frequency prediction. Moreover, it appears that, when increasing the number of time

TABLE III
ERROR FOR DIFFERENT DATASETS (N, VALIDATION SET SIZE, AND n_t, NUMBER OF TIME INSTANTS)

Dataset	Current	Frequency	Time instant
Dataset 1 (N=1,324, n_t=4)	$7.10\ 10^{-2}$	$17.6\ 10^{-2}$	$4.80\ 10^{-2}$
Dataset 2 (N=3,029, n_t=9)	$9.19\ 10^{-2}$	$18.4\ 10^{-2}$	$5.20\ 10^{-2}$
Dataset 3 (N=4,034, n_t=14)	$11.1\ 10^{-2}$	$20.1\ 10^{-2}$	$4.18\ 10^{-2}$
Dataset 4 (N=6,439, n_t=19)	$12.4\ 10^{-2}$	$17.9\ 10^{-2}$	$3.08\ 10^{-2}$
Dataset 5 (N=7,344, n_t=24)	$13.1\ 10^{-2}$	$18.7\ 10^{-2}$	$3.46\ 10^{-2}$

instants, the accuracy of the frequency prediction remains almost unaltered and so does the prediction relevant to the time instant; in contrast, the accuracy of the current prediction decreases.

However, parameters N and n_t proportionally increase: therefore, the information increases, but the number of time instants to learn increases as well. Hence, it is not reasonable to expect a reduction of the error. On top of that, all these training experiments were developed using the same CNN.

V. CONCLUSION

A surrogate model of the inverse induction heating problem has been successfully developed.

In terms of computational costs, the database created for training purpose is based on 1600 multiphysics field analyses in which up to 24 time instants can be identified. For the sake of comparison, an algorithm of evolutionary optimization lasts about a hundred iterations, for a prescribed convergence accuracy, each one requiring an FE analysis. Supposing that the same set of 24 time instants is used, along with the relevant current and frequency pair, the algorithm would approximately require 2400 field analyses, i.e., more than the network training.

In general, it could be stated that once a surrogate model is available, it can be used in optimization and design to great effect and significantly reduce the time to get a solution.

REFERENCES

[1] A. Khan, V. Ghorbanian, and D. Lowther, "Deep learning for magnetic field estimation," *IEEE Trans. Magn.*, vol. 55, no. 6, pp. 1–4, Jun. 2019.
[2] S. Barmada, N. Fontana, L. Sani, D. Thomopulos, and M. Tucci, "Deep learning and reduced models for fast optimization in electromagnetics," *IEEE Trans. Magn.*, vol. 56, no. 3, pp. 1–4, Mar. 2020.
[3] R. Gong and Z. Tang, "Training sample selection strategy applied to CNN in magneto-thermal coupled analysis," *IEEE Trans. Magn.*, vol. 57, no. 6, pp. 1–4, Jun. 2021.
[4] M. Baldan, P. D. Barba, and B. Nacke, "Magnetic properties identification by using a bi-objective optimal multi-fidelity neural network," *IEEE Trans. Magn.*, vol. 57, no. 6, pp. 1–4, Jun. 2021.
[5] *TEAM Problems*. [Online]. Available: https://www.compumag.org/wp/team
[6] P. Di Barba et al., "A benchmark problem of induction heating analysis," *Int. J. Appl. Electromagn. Mech.*, vol. 53, pp. S139–S149, Feb. 2017.
[7] P. Di Barba, F. Dughiero, M. Forzan, and M. E. Mognaschi, "Multiphysics field analysis in induction heating: A CNN-based surrogate model," in *Proc. COMPUMAG*, Cancún, Mexico, Jan. 2022, pp. 1–2.
[8] V. Rudnev, "Subject-oriented assessment of numerical simulation techniques for induction heating applications," *Int. J. Mater. Product Technol.*, vol. 29, nos. 1–4, pp. 43–51, 2007.
[9] *COMSOL Multiphysics*. [Online]. Available: https://www.comsol.com