Data-Driven Deep Convolutional Neural Networks for Electromagnetic Field Estimation of Transformers

Yifan Chen, Qingxin Yang, Yongjian Li, Hao Zhang, and Changgeng Zhang

***Abstract*—This paper aims to estimate the electromagnetic field distribution in a simplified transformer through two-dimensional (2-D) finite element analysis. Traditional neural networks typically take RGB or grayscale image as inputs, which may not be optimal for dense regression tasks encountered in physics field estimations. To address this, we propose a novel approach that constructs datasets incorporating meaningful physical properties, such as magnetic permeability, conductivity and current excitation matrices, as different channels in the input tensors. Specifically, we compare two datasets: the original dataset (OD) simulating grayscale images and the new dataset (ND) integrating distinct material characteristics. This comparative analysis allows us to investigate the impact of different datasets on deep learning performance. Furthermore, to enhance estimation accuracy, we introduce the U-Resnet model, a hybrid architecture combining ResNet's residual blocks with the U-net structure. By comparing the performances of U-net and U-Resnet, we demonstrate the superiority of the latter. Finally, we propose the Add-RMSE loss function, which mitigates the weakening effect of averaging large error pixels when using MSE as the loss function. This enhancement improves gradient propagation during backpropagation and further enhances prediction accuracy. The effectiveness of our proposed method is validated through comprehensive numerical experiments.**

***Index Terms*—Artificial neural network, electromagnetic fields, FEA.**

I. Introduction

NUMERICAL simulation methods, such as the finite element method (FEM) play a crucial role in the electromagnetic analysis of electrical equipment [1], [2]. Despite their high calculation precision, these methods suffer from drawbacks of high cost, long test cycle, instability caused by mesh quality and redundant computational tasks caused by repetitively varied parameters. In recent years, the rapid advancements in high-performance computing and artificial intelligence (AI) technology have revolutionized various fields, including digital image processing, text and language recognition, and autonomous driving [3]. These advancements offer novel approaches to tackle complex physical problem [4].

Artificial neural networks (ANNs) have emerged as powerful tools for data-driven modeling of highly nonlinear dynamic problems. They have demonstrated success in predicting fluid velocity field [5] and stress field [6], significantly reducing calculation time and costs. In the realm of electromagnetic (EM) field estimation, ANNs have been applied to dense regression problems [7], life estimation [8], and fault diagnosis [9]. However, traditional convolutional neural network (CNN) models often fail to preserve spatial and contextual information, leading to the loss of local and global features. Increasing the number of convolution layers to enhance feature utilization can exacerbate issues such as gradient disappearance explosion [10], [11]. Additionally, CNN models commonly utilize RGB images as input [12]. However, for dense regression tasks, this approach may not be optimal for feature extraction. Hence, the exploration of novel methodologies for constructing dataset features offers fresh perspectives on enhancing the predictive efficacy and generalization capability of CNN models.

In this study, a novel model U-Resnet to tackle the challenge of pixel-level dense regression was established. Lightweight ResNet blocks [10], [11] are incorporated into the U-net model [13], leveraging its ability to retain global feature information while enhancing the feature extraction. Furthermore, to explore their impact of these different dataset construction methods on enhancing model prediction accuracy, we constructed two datasets: one incorporating diverse channel-specific physical information, referred to as the new dataset (ND), and the other derived solely from FEM model grayscale images, referred to as the original dataset (OD). Lastly, we introduce the Add-RMSE loss function to effectively emphasize the importance of large pixel errors in improving the loss function during neural network training. Experimental results validate the efficacy of our approaches.

II. Physical Model of Transformer

The simplified 2-D axial symmetrical model of one phase of the three-phase power transformer is depicted in Fig. 1. The structures of clamps, struts, and fans are disregarded, focusing

Manuscript received 31 January 2024; revised 4 April 2024; accepted 6 April 2024. Date of publication 27 June 2024; date of current version 2 October 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 92066206 and Grant 52277008, and in part by the Funds for Creative Research Groups of Hebei Province under Grant E2020202142. *(Corresponding author: Changgeng Zhang.)*

Yifan Chen, Yongjian Li, Hao Zhang, and Changgeng Zhang are with the State Key Laboratory of EERI, School of Electrical Engineering, Hebei University of Technology, Tianjin 300130, China (e-mail: yifanchen_cyf@163.com; liyongjian@hebut.edu.cn; zh709918079@163.com; changgeng.zhang@hebut.edu.cn).

Qingxin Yang is with the Tianjin University of Technology, Tianjin 300384, China (e-mail: qxyang@tjut.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASC.2024.3420184.

Digital Object Identifier 10.1109/TASC.2024.3420184

1051-8223 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on November 25,2024 at 01:39:58 UTC from IEEE Xplore. Restrictions apply.

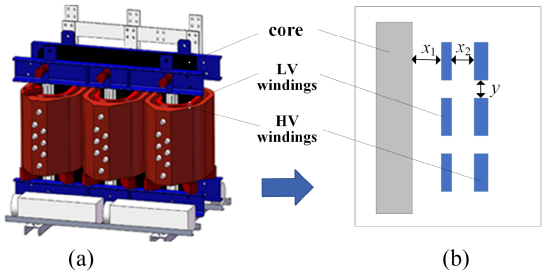


Fig. 1. Physical model. (a) Dry-type transformer. (b) Simplified 2-D axial symmetrical model.

solely on the core and winding components, which play a pivotal role in the electromagnetic field distribution. The governing equations of the eddy current region are as follows:

$$\nabla \times \left(\frac{1}{\mu}\nabla \times \boldsymbol{A}\right) = -\left(\boldsymbol{J}_s + \boldsymbol{J}_e\right) \tag{1}$$

$$\boldsymbol{J}_e = -\sigma_e\left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\varphi\right) \tag{2}$$

$$\nabla \cdot \left(\boldsymbol{J}_s + \boldsymbol{J}_e\right) = 0 \tag{3}$$

where $\boldsymbol{A}$ is the magnetic vector potential, φ is the electric scalar potential, μ is the magnetic permeability tensor, $\boldsymbol{J}_s$ is the source current density of the conductor, $\boldsymbol{J}_e$ is the eddy current density induced by leakage flux, and σ_e is the electric conductivity.

III. CNN Architecture

A. U-Net

The U-net architecture, with its symmetric encoder-decoder structure comprising 4 encoder layers, 4 decoder layers, and skip-connections to concatenate feature maps from shallower and deeper layers, integrates low-level information for precise localization and semantic detail extraction [13]. Hence, U-net is selected as the baseline for EM field estimation, given its adaptable structure and robust feature extraction capabilities.

B. ResNet

To enhance feature extraction capabilities, the layers of the deep neural network (DNN) are progressively deepened. However, this may lead to issues such as vanishing or exploding gradients, resulting in ineffective updates of deep weight parameters. To address this, shortcut connections between a few stacked layers are introduced as Fig. 2. These shortcut connections transfer the approximation of a complex function $H(x)$ to a residual function $F(x) = H(x) - x$, enabling deeper networks to perform no worse than shallower ones and improving sensitivity to minute changes [10]. The purity of the identity mapping and the sequence of activation functions in the residual path are found to impact the network's performance [11]. The residual unit structure depicted in Fig. 3 exhibits optimal performance and is implemented in the ResNet models.

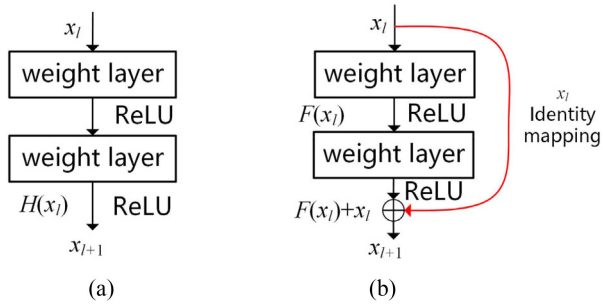


Fig. 2. The block of feature information between DNN layers. (a) Original block. (b) Residual block with identity mapping in ResNet.

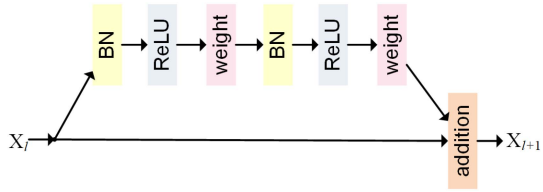


Fig. 3. Full pre-activation residual unit.

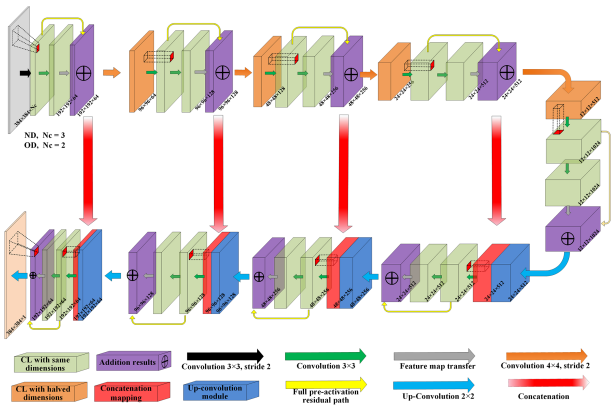


Fig. 4. The overview architecture of U-ResNet.

C. U-ResNet Architecture

To further enhance the feature extraction capability of U-net in dense regression tasks, we have integrated ResNet residual blocks instead of conventional convolutional layers (CL), resulting in the construction of the U-Resnet model, as illustrated in Fig. 4. This model consists of four down-sampling modules and four up-sampling modules, each incorporating a shortcut connection and a residual path. The residual path consists of two CLs pre-activated by batch normalization and ReLU layers, as depicted in Fig. 3, along with an additional layer for the element-wise addition operation. The predictive performance of two models, U-net and U-Resnet, on different datasets are compared in Section IV-B.

IV. Magnetic Field Estimation

A. Data Preprocessing

The quality and scale of the dataset significantly impact the training of DNN weight parameters. While U-net performs well

Authorized licensed use limited to: Hebei University of Technology. Downloaded on November 25,2024 at 01:39:58 UTC from IEEE Xplore. Restrictions apply.

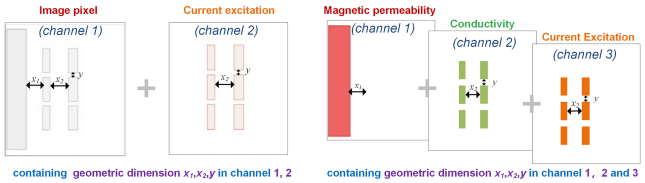


Fig. 5. Two types of datasets. (a) Original grayscale images dataset. (b) New dataset.

TABLE I
PARAMETERS OF THE PREPROCESSING IMAGES DATASET

Physical quantities	Parameters	Actual value	Normalized tensor value
Geometric dimension[mm]	x_1	10-70	--
	x_2	20-80	--
	y	15-75	--
Current excitation[A]	I_{max}	2000-2800	0-1
Relative Permeability	μ_{steel}	8000	1
	μ_{copper}	1	0
	μ_{air}	1	0

with small datasets, high-quality datasets are essential to prevent overfitting and ensure effective weight parameter training. Leveraging Maxwell's parameterized modeling, a simplified 2-D finite element model of a single-phase dry-type transformer is established. The geometric model, along with its computed magnetic density distribution, serves as input X and label Y, respectively.

To investigate the impact of different datasets construction methods on enhancing model prediction accuracy, two datasets are constructed: one incorporating diverse channel-specific physical information, referred to as the new dataset (ND), and the other derived from FEM geometric model images, referred to as the original dataset (OD). Specifically the ND includes material characteristics namely magnetic permeability, conductivity and current excitation. Conversely, the OD simulates grayscale images of the transformer model along with current excitation. Both ND and OD incorporate geometric dimensions parameters including the horizontal distances x_1, x_2 and vertical distance y depicted in Fig. 5. Here, x_1 and x_2 represent horizontal distances between the core and low-voltage (LV) windings and between LV and high-voltage (HV) windings, respectively. y indicates the vertical spacing between layers in segmented windings, with their variation ranges listed in Table I.

To enhance the generalization ability, the training set was augmented by flipping operations. Each dataset comprises 8010 samples, with 7600 allocated to the training set, 400 to the validation set, and 10 to the test set.

B. Experimental Results of Magnetic Field Estimation

During the training phase, we compared prediction accuracies among different model architectures, datasets, numbers of convolutional kernels, and loss functions. Table II provides detailed information on specific hyperparameters and model accuracy. Our models were trained using the TensorFlow 2.9.1 framework on an NVIDIA A40 GPU with 40 GB of video memory. A fixed learning rate of 0.001 was applied across all models, with a maximum of 200 training epochs.

TABLE II
MODEL ERRORS ACROSS DIFFERENT HYPERPARAMETERS AND DATASETS

CNN models	Datasets	Loss function	Kernel number times	Model size	Error (MAE)
U-net	ND	MSE	1	4M	8.56%
	ND	ADD-RMSE	1	4M	7.21%
	ND	MSE	2	16M	5.99%
	ND	ADD-RMSE	2	16M	4.59%
	OD	MSE	1	4M	12.41%
	OD	ADD-RMSE	1	4M	11.63%
	OD	MSE	2	16M	9.78%
	OD	ADD-RMSE	2	16M	8.86%
U-Resnet	ND	MSE	1	4M	2.87%
	ND	ADD RMSE	1	4M	2.51%
	ND	MSE	2	16M	1.78%
	ND	**ADD-RMSE**	**2**	**16M**	**1.56%**
	OD	MSE	1	4M	3.59%
	OD	ADD-RMSE	1	4M	3.11%
	OD	MSE	2	16M	2.35%
	OD	ADD-RMSE	2	16M	1.93%

In traditional neural network regression tasks, Mean Squared Error (MSE) and Mean Absolute Error (MAE) are commonly used loss functions. However, predicting magnetic fields presents greater complexity than typical image outputs. In our dense regression problems, the output is a matrix of 384 $*$ 384 floating-point numbers, requiring higher precision and a broader range than the integer values (0–255) typically used in images. Using MSE or MAE would average errors across all pixels, potentially obscuring errors at individual pixel points and impeding model training. To mitigate this issue, we introduce the Add_RMSE loss function defined by (4).

$$Loss_{Add_RMSE} = \sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{(y_{i,j}-\hat{y}_{i,j})^2} \tag{4}$$

Where, $y_{i,j}$ denotes the magnetic field value (in pixels) at the ith row and jth column of the ground truth, while $\hat{y}_{i,j}$ represents the predicted magnetic field value (in pixels) at the same position in the neural network output. The loss function Add_RMSE is designed to emphasize the importance of large error pixels for improving the loss function during neural network training. This approach ensures that large error pixels have a stronger influence on parameter updates during gradient computation, thereby improving the training effectiveness and prediction accuracy of the model.

Specifically, Add_RMSE first calculates the element-wise difference between the model predictions and the ground truth values, then computes the L2 norm of these differences, and finally sums the norms across all pixels. The purpose is to strengthen the response of the loss function to large error pixels. In this way, during backpropagation, gradients can propagate better to these pixels, allowing the model to focus more on these important regions and thereby improving prediction accuracy.

As shown in Table II, the U-Resnet model with residual blocks demonstrates significantly higher prediction accuracy compared to the U-net model. Additionally, ND outperforms OD due to its distinct physical interpretation of material properties and easily identifiable features. However, ND has its limitations.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on November 25,2024 at 01:39:58 UTC from IEEE Xplore. Restrictions apply.

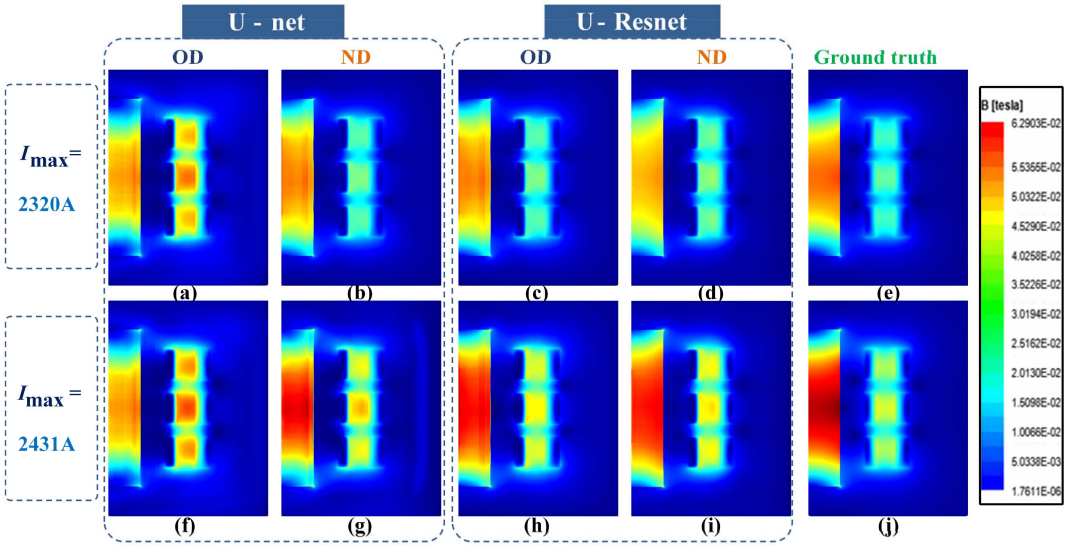


Fig. 6. Magnetic field distribution estimation through U-net and U-ResNet.

One observed during training is the increased volatility of errors, making overfitting more likely. This could be due to the excessive number of zero values in ND, which hampers effective parameter training in the neural network.

The proposed Add-RMSE loss function effectively highlights pixels with higher error magnitudes, guiding the network's weight parameters to optimize and reduce errors at these points. Moreover, the increased utilization of convolution kernels enhances the encoder layers' ability to extract high-dimensional features, facilitating smoother electromagnetic field distribution data generation in the subsequent decoder layers. Nonetheless, it's essential to note that more convolution kernels don't guarantee better performance. Excessive kernel usage demands additional computational resources and may lead to overfitting.

Fig. 6 illustrates the prediction results of the neural network utilizing the Add-RMSE loss function and a kernel numbers factor of 2. As observed, the U-Resnet model exhibits notably superior prediction results compared to the U-net model when applied to the ND. Specifically, under the same excitation conditions and dataset the U-ResNet (UR) model demonstrates smoother transitions and higher accuracy compared to the traditional U-Net (U) model. The improvement is evident in Fig. 6. At I = 2320 A, the performance of UR under OD (Fig. 6(c)) is closer to the ground truth (Fig. 6(e)) compared to U's results (Fig. 6(a)). Similarly, the UR model under ND (Fig. 6(d)) exhibits closer alignment with the ground truth (Fig. 6(e)) compared to U (Fig. 6(b)). Likewise, at I = 2431 A, the performance of UR under OD (Fig. 6(h)) better matches the ground truth (Fig. 6(j)) compared to U (Fig. 6(f)), while the UR model under ND (Fig. 6(i)) displays closer proximity to the ground truth (Fig. 6(j)) compared to U (Fig. 6(g)). These discrepancies primarily stem from variations in the magnetic field values and smoothness around the core and windings regions. Nevertheless, it is important to notice that some errors still exist between the predicted results and the ground truth. This could potentially be attributed to the limited size of the dataset. Future endeavors will focus on expanding the dataset size to enhance the accuracy of predictions.

V. Conclusion

In this study, we have implemented a CNN model to estimate the EM field distribution in electrical equipment, significant for real-time applications like digital twins, enhancing the operational efficiency and safety of electrical equipment. Several significant conclusions are drawn:

1) The introduction of the new dataset (ND), which integrates physical information such as material properties, current excitations, and geometric dimensions, has demonstrated superior effectiveness in enhancing precision compared to the original dataset (OD). This approach facilitates more efficient feature extraction, thereby enhancing prediction precision. However, it is noteworthy that the prevalence of numerous zero values in the ND poses challenges, as most convolutional kernels struggle to receive effective training, resulting in slower convergence of neural networks trained on ND.
2) Our experimental results unequivocally show that the U-Resnet architecture surpasses the traditional U-net model. Capitalizing on its residual connections, U-Resnet achieves more accurate estimations on the regions around core and windings, exhibiting smoother results despite some residual errors. This highlights the effectiveness of residual connections in mitigating the impact of errors and improving prediction accuracy.
3) We introduce the Add-RMSE loss function, which distinguishes itself from the traditional MSE loss function by substituting the averaging operation with addition. Add-RMSE prioritizes the contribution of large error pixels to the loss, facilitating more effective parameter updates during model training. By capturing cumulative errors associated with large error points, Add-RMSE enhances accuracy for dense regression problems.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on November 25,2024 at 01:39:58 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Y. Li, X. Yan, C. Wang, Q. Yang, and C. Zhang, "Eddy current loss effect in foil winding of transformer based on magneto-fluid-thermal simulation," *IEEE Trans. Magn.*, vol. 55, no. 7, Jul. 2019, Art. no. 8401705.

[2] Y. Chen, Q. Yang, C. Zhang, Y. Li, and J. Wang, "Magnetic-thermal coupling analysis of saturable reactor," *IEEE Trans. Magn.*, vol. 58, no. 2, Feb. 2022, Art. no. 6300705.

[3] M. Guo et al., "Attention mechanisms in computer vision: A survey," *Comput. Vis. Media*, vol. 8, no. 3, pp. 331–368, 2022.

[4] S. Rudy et al., "Data-driven discovery of partial differential equations," *Sci. Adv.*, vol. 3, no. 4, 2017.

[5] J. Ling., A. Kurzawski, and J. Templeton, "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance," *J. Fluid Mech.*, vol. 807, pp. 155–166, 2016.

[6] Z. Nie, H. Jiang, and L. B. Kara, "Stress field prediction in cantilevered structures using convolutional neural networks," *J. Comput. Inf. Sci. Eng.*, vol. 20, no. 1, 2020, Art. no. 011002.

[7] A. Khan, V. Ghorbanian, and D. Lowther, "Deep learning for magnetic field estimation," *IEEE Trans. Magn.*, vol. 55, no. 6, Jun. 2019, Art. no. 7202304.

[8] S. Forouhari and A. Abu-Siada, "Application of adaptive neuro fuzzy inference system to support power transformer life estimation and asset management decision," *IEEE Trans. Dielectrics Elect. Insul.*, vol. 25, no. 3, pp. 845–852, Jun. 2018.

[9] Z. Li et al., "Fault diagnosis of transformer windings based on decision tree and fully connected neural network," *Energies*, vol. 14, no. 6, 2021, Art. no. 1531.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[11] K. He et al., "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 630–645.

[12] R. Gong et al., "Investigation of convolutional neural network U-net under small datasets in transformer magneto-thermal coupled analysis," *COMPEL- Int. J. Computation Math. Elect. Electron. Eng.*, vol. 39, no. 4, pp. 959–970, 2020.

[13] O. Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on November 25,2024 at 01:39:58 UTC from IEEE Xplore. Restrictions apply.