Multi-Channel Neural Networks-Based Thermal Monitoring of Electric Motor

Jaehoon Shim, Jonghun Choi, Sangwon Lee and Jung-Ik Ha
Department of Electrical and Computer Engineering Seoul National University, Seoul, Korea

***Abstract*--This paper presents a thermal monitoring model for electric drive motors using Multi-Channel Neural Networks (MCNNs). The proposed structure consists of three key components. Firstly, a data-driven approach is adopted, eliminating the need for additional sensors, except during the training phase. Secondly, it employs a multi-channel architecture that processes two sets of inputs. One channel handles input power-related thermal modeling, while the other focuses on modeling thermal losses. Consequently, these channels are combined to create a data-driven thermal system. Thirdly, a suitable structure is proposed for the input data, tailored differently for each channel. Furthermore, the MCNNs are trained end-to-end using the provided loss function. The model's performance is assessed using data collected from an 89 kW induction motor. Additionally, comparative models based on previously studied machine learning-based thermal models are utilized for comparison.**

***Index Terms*--Thermal monitoring, electric motor, Multi-Channel Neural Networks (MCNNs), Data-driven modeling.**

I. Introduction

The importance of monitoring electric power systems has been emphasized due to the increasing prevalence of electric power-based transportation methods. As a result, numerous studies have been conducted to diagnose various aspects, including the diagnosis of inverters in electric drive systems [1], mechanical fault diagnosis [2], and research on thermal monitoring [3]. This study concentrates explicitly on thermal monitoring. With the rapid expansion and high performance of electric vehicles, the significance of thermal monitoring for electric motors has grown. It plays a crucial role in ensuring safety and preventing motor damage caused by the demagnetization of permanent magnets.

Various studies have been conducted for thermal monitoring, and the extensively researched methods are indirect methods that estimate the temperatures by observing electrical parameters related to temperatures [3]. These methods are further divided into signal injection-based methods [4], [5] and non-invasive-based methods [4], [6], [7] (i.e., without signal injection). Injecting signals has the advantage of being able to estimate temperatures regardless of driving conditions by observing the parameters of high-frequency components. However, it introduces noise issues and leads to unwanted torque ripple caused by injected signals. Furthermore, it can evoke deterioration of efficiency.

Regarding the non-invasive methods, the parameter estimations for temperature estimation are considered representative. The resistor of the stator winding is utilized as it relies on the temperatures of the copper. At the same time, the flux density of the permanent magnet is also available for the same purpose. However, estimating these parameters based on voltage equations introduces numerous uncertainties. Additionally, observation difficulties arise under specific driving conditions.

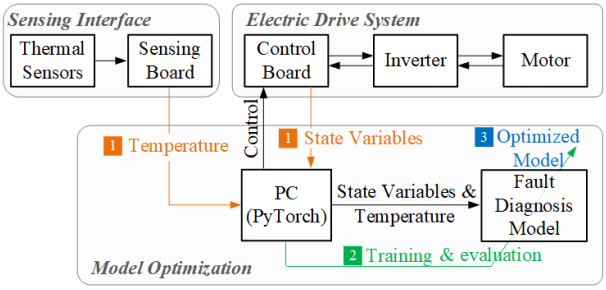
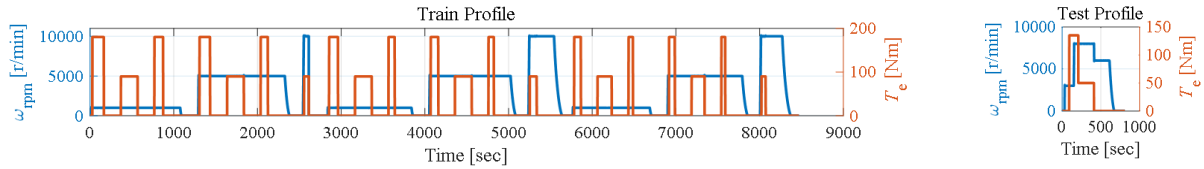


Fig. 1. Steps for developing the data-driven thermal monitoring model.

Recently, data-based methods utilizing Neural Networks (NNs) have been widely studied to address these issues [8]–[10]. [11] proposes estimating temperatures by integrating the temperature variations between the current and following sampling points. However, this method requires additional integration outside of the NNs, which is cumbersome. Additionally, there is a design inefficiency in that domain knowledge is not reflected when constructing the NNs. In [8], a machine learning-based lumped-parameter thermal model is presented. The model effectively incorporates domain knowledge. However, the use of explicit lumped-parameter thermal networks restricts the flexibility of the thermal model. Therefore, this paper proposes the MCNNs-based thermal monitoring method using data-driven approaches to overcome the abovementioned limitations.

This paper utilizes a combined structure of difference-estimating feedforward neural network (DFNN) and vanilla feedforward neural network (FNN) among various machine learning models. A comprehensive comparative analysis is conducted to establish the superiority of the proposed model architecture. It includes the thermal monitoring model based on recurrent neural network (RNN), FNN, and DFNN.

Several steps are involved in developing the data-driven model, as depicted in Fig. 1. The initial phase consists in gathering data from the sensing and control board, represented in orange. This information includes currents, voltage, torque, stator temperature, winding temperature, and speed. These data are then used for training the thermal



(a) (b)

Fig. 2. Operational profile for data acquisition. (a) For the training dataset. (b) For the testing dataset.

TABLE I
DEFINITION OF DATA

Data Structure	Variables	Symbols
Power related data ($\boldsymbol{U} \in \mathbb{R}^{7\times1}$)	d-axis voltage reference	$v^r_{ds,ref}$
	q-axis voltage reference	$v^r_{qs,ref}$
	d-axis current	i^r_{ds}
	q-axis current	i^r_{qs}
	Slip freqency	ω_{sl}
	Mechanical angular velocity	ω_{rpm}
	Temperature of housing cooling water	T_{env}
Temperature data by parts ($\boldsymbol{T} \in \mathbb{R}^{5\times1}$)	Temperature of stator winding	$T_{winding}$
	Temperature of jumper	T_{jump}
	Temperature of neutral point	$T_{neutral}$
	Temperature of housing	$T_{housing}$
	Temperature of rotor	T_{rotor}

model. The training and evaluation processes are conducted using the PyTorch deep-learning library on a PC in this study. These processes are visually depicted using green and blue colors.

The main contributions of this research are as follows:

1) The MCNNs-based model is data-driven and does not require additional sensors. Instead, it utilizes signals already accessible in the motor controller, making it cost-effective and practical.
2) The proposed structure consists of two channels of neural networks (NNs) that effectively describe the thermal physical system. One channel focuses on the input power-related thermal model, while the other describes the thermal loss-related model. It allows for a comprehensive representation of the system.
3) An objective function is proposed and designed for the architecture to enable end-to-end learning. In addition, experimental verification is carried out using an 89 kW real-world induction motor.

The remainder of this paper is organized as follows: Section II presents the data acquisition strategy and the method for dividing the data into training and evaluation sets. In Section III, the proposed model architecture and learning method are presented. Section IV presents the evaluation results based on the experimental data. Moreover, Section V provides the conclusions.

II. DATA ACQUISITION STRATIGIES

This section emphasizes two main topics. Firstly, it examines the data types acquired and utilized as input data. Secondly, the process of obtaining these data while driving the motor is discussed.

A. Data Types for Temperature Estimation

Table I presents two types of data structures: power-related variables and temperature data. Power-related data are utilized to estimate temperature variations for the subsequent step. On the other hand, temperature data simultaneously serve as the model's ground truth and input values.

Regarding the power-related data, $\boldsymbol{U} \in \mathbb{R}^{7\times1}$, it comprises variables associated with the output power of the target motor. Specifically, this study considers the following power-related variables: $v^r_{ds,ref}$, $v^r_{qs,ref}$, i^r_{ds}, i^r_{qs}, ω_{sl}, and ω_{rpm}, and the temperature of the housing cooling water, T_{env}. This research considers the cooling water surrounding the motor system as the environment.

Concerning the temperatures data, $\boldsymbol{T} \in \mathbb{R}^{5\times1}$, it encompasses measurements from five distinct locations within the motor: winding, jumper, neutral, housing, and rotor. These locations are presented in Table I.

B. Driving Profile for Data Acquisition

The data-gathering method plays a crucial role in the success of supervised learning. Collecting comprehensive and diverse data is essential to build a reliable data-driven model. Therefore, careful consideration is given to selecting operational points where data is collected, covering a wide range of torque and velocity profiles. The collected data, including various torque and velocity combinations, is presented in Fig. 2. Additionally, to ensure the versatility of the trained model, separate profiles are designed for the training dataset, as shown in Fig. 2(a), and the test dataset as shown in Fig. 2(b). This approach guarantees the model can generalize well to new data beyond the training set.

II. PROPOSED MULTI-CHANNEL-BASED MODEL

This section covers three main points. Firstly, it introduces the proposed input data structure based on the predefined data structure in Table I. Secondly, it presents the architecture of the proposed model, which consists of two channels of neural networks. Lastly, it discusses the optimization method designed for end-to-end learning.

A. Input Data Configuration

The input data is constructed using the $\boldsymbol{U}$ and $\boldsymbol{T}$ vectors, considering the time series characteristic of the

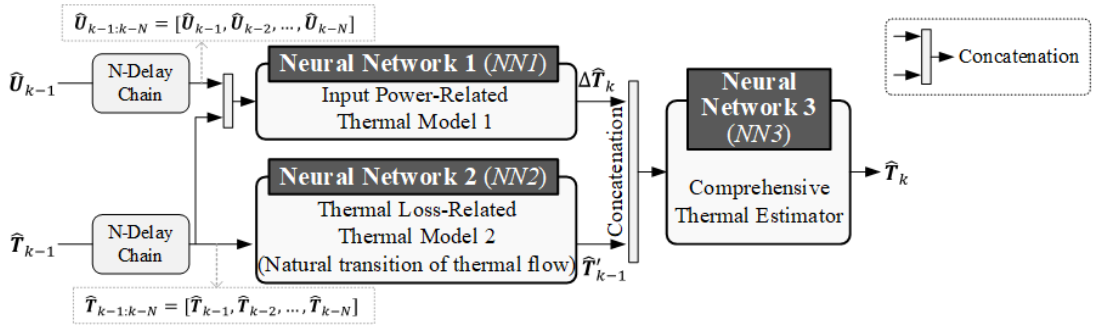


Fig. 3. The proposed model structure. It comprises three NNs, each with a distinct role. One NN is responsible for the input power-related thermal model, while another NN estimates temperatures for the next time step, considering thermal loss. The third NN serves as a comprehensive thermal estimator.

temperature estimation. The input data is defined with a corresponding time stamp to incorporate the time aspect. The basic definitions of these concepts are represented in (1) and (2):

$$\boldsymbol{U}_k = \{|\left(V^r_{ds,ref}\right)_k|, |\left(V^r_{qs,ref}\right)_k|, |\left(I^r_{ds,ref}\right)_k|, |\left(I^r_{qs,ref}\right)_k|, \ |(\omega_{sl})_k|, |\left(\omega_{rpm}\right)_k|\} \quad (1)$$

$$\boldsymbol{T}_k = \{T_{winding,k}, T_{jump,k}, T_{neutral,k}, T_{housing,k}, T_{rotor,k}\} \quad (2)$$

where k represents the kth measured data structure in the time domain.

Using (1) and (2), the final structures are built as an N-delay chain. In other words, this means that the frames are formed with a window size of N samples. These structures are represented by (3) and (4):

$$\boldsymbol{U}_{k-1:k-N} = [\boldsymbol{U}_{k-1}, \boldsymbol{U}_{k-2}, \dots, \boldsymbol{U}_{k-N}] \quad (3)$$

$$\hat{\boldsymbol{T}}_{k-1:k-N} = [\hat{\boldsymbol{T}}_{k-1}, \hat{\boldsymbol{T}}_{k-2}, \dots, \hat{\boldsymbol{T}}_{k-N}] \quad (4)$$

Here, $\hat{\cdot}$ indicates estimated values. Noticeably, the $\boldsymbol{T}$ is the output at the current time and serves as the input data for the next time step.

B. Model Architecture

The proposed model's main idea is overcoming thermal monitoring difficulties via linear system modeling. The existing linear models are represented as (5)-(7):

$$\Delta\boldsymbol{T}_k = \boldsymbol{A} \times \boldsymbol{T}_{k-1} + \boldsymbol{B} \times \boldsymbol{U}_{k-1} \quad (5)$$

$$\boldsymbol{T}_k = \boldsymbol{C} \times \boldsymbol{T}_{k-1} \quad (6)$$

$$\boldsymbol{T}_k = \boldsymbol{T}_{k-1} + \Delta\boldsymbol{T}_k \quad (7)$$

where $\boldsymbol{A}$ represents the change in temperatures considering the motor power given the current motor temperatures. $\boldsymbol{B}$ consists of input power-related matrices that describe how motor power influences the temperature fluctuation in the motor. $\boldsymbol{C}$ is attributed to natural state transition, which is predetermined by the design of the motor. These matrices are difficult to determine accurately. Even with the best determination, the linear model has limitations in capturing the nonlinearity of the thermal model.

The proposed architecture of the model where NNs replace each matrix is proposed to overcome the limitations of linear system-based modeling as described above. In other words, the proposed model transforms the linear model into a nonlinear model using multiple NNs.

The proposed structure is illustrated in Fig. 3. *NN1* is responsible for estimating the temperature variations based on the input vectors, $\boldsymbol{U}_{k-1:k-N}$ and $\hat{\boldsymbol{T}}_{k-1:k-N}$, similar to the $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices in (5). Using previous temperatures, *NN2* estimates the temperatures that naturally transition to the next state. Finally, *NN3* combines the outputs of *NN1* and *NN2* to provide a comprehensive thermal estimation, similar to the approach described in (7). The relationships between the inputs and outputs of each module are illustrated in (8)-(10):

$$\Delta\hat{\boldsymbol{T}}_k = f^{NN1}\left(\mathcal{C}\left((\boldsymbol{U}_{k-1:k-N})_n, \left(\hat{\boldsymbol{T}}_{k-1:k-N}\right)_n\right)\right) \quad (8)$$

$$\hat{\boldsymbol{T}}'_k = f^{NN2}\left(\left(\hat{\boldsymbol{T}}_{k-1:k-N}\right)_n\right) \quad (9)$$

$$\hat{\boldsymbol{T}}_k = f^{NN3}\left(\mathcal{C}\left(\left(\hat{\boldsymbol{T}}'_k\right)_n, \left(\Delta\hat{\boldsymbol{T}}_k\right)_n\right)\right) \quad (10)$$

where $\mathcal{C}(\cdot, \cdot)$ indicates the concatenation of the values. It is worth noting that the results of (8) and (9) are concatenated and used as input to NN3 for estimating the temperatures vector, $\hat{\boldsymbol{T}}_k$, for the next time step, k.

The basic calculations of the proposed three NNs are shown in (11) and (12):

$$\zeta^{l^*}(\cdot; \theta^{l^*}) = \begin{cases} ReLU\left(\left(W^{l^*}\right)^T(\cdot) + b^{l^*}\right), & l^* < l^*_d \\ \left(W^{l^*}\right)^T(\cdot) + b^{l^*}, & l^* = l^*_d \end{cases}$$

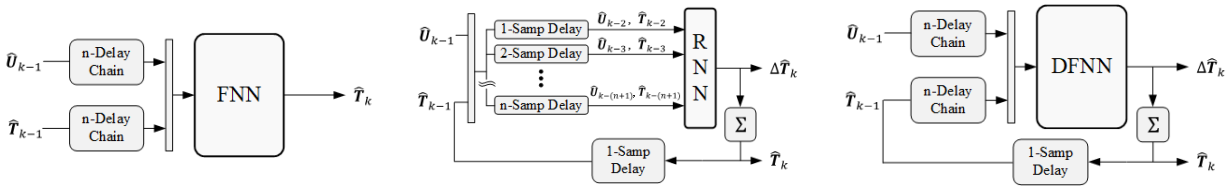
$$\theta^{l^*} = \{W^{l^*}, b^{l^*}\}, \quad *\in \{NN1, NN2, NN3\} \quad (11)$$

$$f^*(\cdot; \boldsymbol{\Theta}^{l^*}) = \zeta^*_{l^*_d}(\ ; \theta^{l^*_d}) \circ \cdots \circ \zeta^*_1(\cdot; \theta^{l^*_1}) \quad (12)$$

where W^{l^*} and b^{l^*} are weights and biases of l^*th layer, and θ^{l^*} is the set of l^*th layer's parameters. l^*_d is the total number of layers of $*$ network. $\boldsymbol{\Theta}^{l^*}$ is all of the parameters of the $*$ network. In addition, $\circ$ indicates the serial calculation of the layers of NNs.

C. Model Optimization

The optimization of the proposed model involves multiple NNs, but it is designed to be trained simultaneously through end-to-end learning using a single



(a) (b) (c)

Fig. 4. Comparison models. (a) Feedforward neural network (FNN)-based model, (b) Recurrent neural network-based model, and (c) Difference estimating feedforward neural network model.

TABLE III
EVALUATION RESULTS

Model \ MSE	MSE of $T_{winding}$		MSE of T_{jump}		MSE of $T_{neutral}$		MSE of $T_{housing}$		MSE of T_{rotor}	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
FNN-Based Model	13.92	4.76	9.49	2.76	12.37	3.42	6.66	1.83	14.47	4.37
RNN-Based Model	9.96	3.36	8.56	2.21	8.88	3.11	8.38	5.92	13.82	4.68
DFNN-Based Model	5.31	0.98	4.22	1.13	4.76	2.22	3.77	0.42	5.46	2.49
Proposed Model	1.16	0.39	1.01	0.52	0.58	0.24	0.39	0.15	1.31	0.56

TABLE II
PARAMETERS USED IN PROPOSED MODEL OPTIMIZATION

Hyperparemeters	Value
Epochs	200
Batch Size	512
Learning rate (lr)	0.001
Optimization	Adam
Data preprocess	Min-max scaling (range: -1~1)
Order of N-delay	3
Number of Nodes in Hidden Layer	20
Depth of Hidden layer	3

objective function. This optimization rule is represented in (13):

$$\left\{\boldsymbol{\Theta}^{l^{NN1}}, \boldsymbol{\Theta}^{l^{NN2}}, \boldsymbol{\Theta}^{l^{NN3}}\right\}^{Opt} = \underset{\{\boldsymbol{\Theta}^{l^{NN1}}, \boldsymbol{\Theta}^{l^{NN2}}, \boldsymbol{\Theta}^{l^{NN3}}\}}{\operatorname{argmin}} (\mathcal{L}) \quad (13)$$

The proposed objective function ($\mathcal{L}$) is formulated with two terms to achieve it. In addition, each term is designed to ensure that each NN fulfills its intended purpose, as mentioned in subsection *B*. The proposed objective function is depicted in (14):

$$\mathcal{L} = \mathcal{L}_{diff} + \mathcal{L}_{est} \quad (14)$$

It comprises two loss terms: $\mathcal{L}_{diff}$ for considering the f^{NN1} channel and $\mathcal{L}_{est}$ for considering the f^{NN2} channel. Although the loss function for *NN3* is not explicitly proposed, comprehensive optimization of *NN3* is achieved through (14). These loss terms are respectively defined in (15) and (16):

$$\mathcal{L}_{diff} = \frac{1}{N_B}\sum_{i=1}^{N_B}\left((\Delta \boldsymbol{T}_k)_n - \left(\Delta \widehat{\boldsymbol{T}}_k\right)_n\right)^2 \quad (15)$$

$$\mathcal{L}_{est} = \frac{1}{N_B}\sum_{i=1}^{N_B}\left((\boldsymbol{T}_k)_n - \left(\widehat{\boldsymbol{T}}_k\right)_n\right)^2 \quad (16)$$

where N_B is the batch size that is used for training the model. Each value represents the mean squared error (MSE) between the ground truth and the estimated.

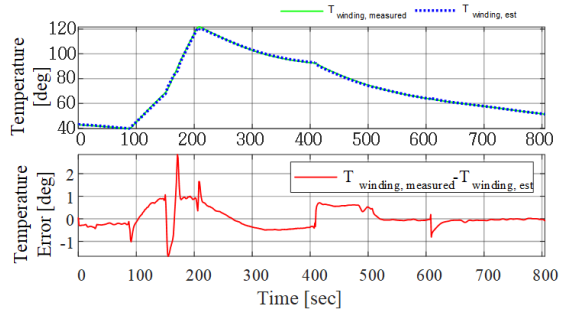
The hyperparameters used for model optimization are listed in Table II.

IV. EXPERIMENTAL RESULTS

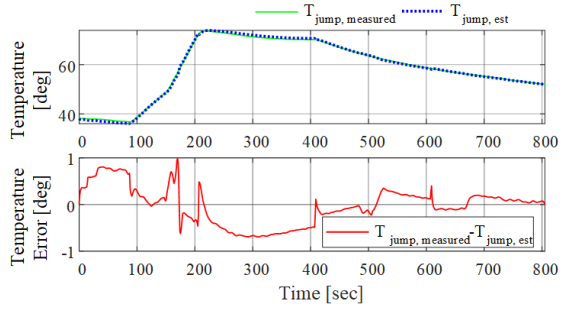
The proposed model is compared with recently developed machine learning-based temperature estimation models, as illustrated in Fig. 4. The selection of the comparison target model is based on the [11]. Fig. 4(a) presents the vanilla FNN-based temperatures estimation model. In this structure, the temperatures of the following state are estimated using the previous state's temperatures and the current motor output information. Fig. 4(b) depicts the RNN structure, where the temperature variations are evaluated, and the next-step temperatures are determined by adding the temperature variations with an external integrator. Fig. 4(c) represents the structure proposed in [11], where the FNN structure is employed instead of the RNN in Fig. 4(b) to estimate the variations in the temperature.

The results of the comparison between the proposed model and the benchmark models are presented in Table III. For the evaluation, the same dataset is used for all the models. The proposed model shows superior performance compared to the benchmark models in a quantitative manner. The results of the comparison between the proposed model and the benchmark models are presented numerically in Table III. The evaluation is conducted using the same dataset for all models, and the evaluation metric is the mean squared error (MSE). The proposed model demonstrates superior performance compared to the benchmark models in a quantitative manner. Compared to the DFNN-based model, the proposed model achieves reductions in MSE of 78.15 %p (i.e., 5.31 to 1.16) for $T_{winding}$, 76.06 %p (i.e., 4.22 to 1.01) for T_{jump}, 87.82 %p (i.e., 4.76 to 0.58) for $T_{neutral}$, 89.66 %p (i.e., 3.77 to 0.39) for $T_{housing}$, and 76.01 %p (i.e., 5.46 to 1.31) for T_{rotor}.

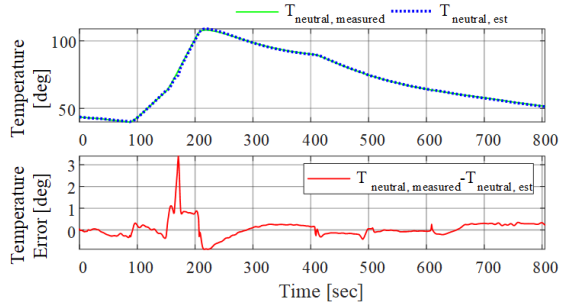
The test results in Fig. 5 visually compare the ground truth and estimated temperatures. The green line represents the ground truth, while the blue dot line represents the estimated values. The difference between them, shown by the red line, remains within 5 degrees. It indicates a relatively small error percentage, especially considering the temperatures exceed 100 degrees.



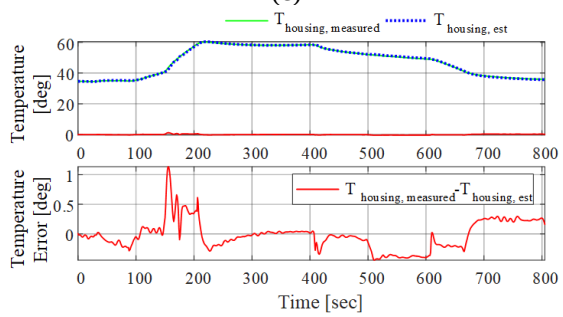
(a)



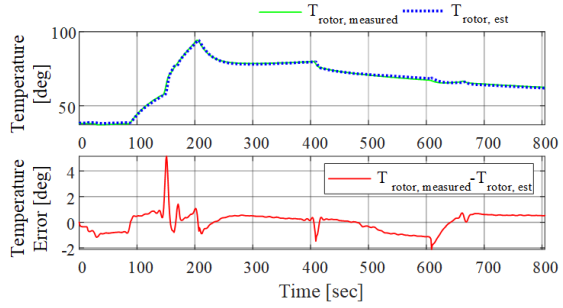
(b)



(c)



(d)



(e)

Fig. 5. Test results graph depicting the measured and estimated values for the following components: (a) winding, (b) jumper, (c) neutral point, (d) housing, and (e) rotor.

V. Conclusions

This paper presents a thermal monitoring model based on neural networks with a multi-channel architecture. The model consists of three neural networks that capture the input power-related thermal behavior, natural transition of thermal flow, and estimate temperatures for the next step. A suitable loss function is proposed to ensure practical training through end-to-end learning. Experimental results using real-world data from an 89 kW induction motor demonstrate the superior performance of the proposed model compared to existing models in accurately estimating temperatures.

References

[1] J. Shim, J. Lee, and J.-I. Ha, "Current-Sensor and Switch-Open Fault Diagnosis Based on Discriminative Machine Learning Model for PMSM Driving System," in 2021 IEEE Energy Conversion Congress and Exposition (ECCE), Vancouver, BC, Canada: IEEE, Oct. 2021, pp. 5098–5104. doi: 10.1109/ECCE47101.2021.9595123.

[2] J. Shim, T. Joung, S. Lee, and J.-I. Ha, "Audio Data-driven Anomaly Detection for Induction Motor Based on Generative Adversarial Networks," in 2022 IEEE Energy Conversion Congress and Exposition (ECCE), Detroit, MI, USA: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/ECCE50734.2022.9947652.

[3] O. Wallscheid, "Thermal Monitoring of Electric Motors: State-of-the-Art Review and Future Challenges," IEEE Open J. Ind. Applicat., vol. 2, pp. 204–223, 2021, doi: 10.1109/OJIA.2021.3091870.

[4] H.-S. Jung, H. Kim, S.-K. Sul, and D. J. Berry, "Magnet Temperature Estimation of Traction Motor in Standstill With Considering Spatial Harmonics," IEEE Trans. Ind. Electron., vol. 68, no. 11, pp. 10546–10557, Nov. 2021, doi: 10.1109/TIE.2020.3031516.

[5] H. Kim, H.-S. Jung, and S.-K. Sul, "Stator Winding Temperature and Magnet Temperature Estimation of IPMSM Based on High-Frequency Voltage Signal Injection," IEEE Trans. Ind. Electron., vol. 70, no. 3, pp. 2296–2306, Mar. 2023, doi: 10.1109/TIE.2022.3174285.

[6] H.-S. Jung, H. Kim, S.-K. Sul, and D. J. Berry, "Temperature Estimation of IPMSM by Using Fundamental Reactive Energy Considering Variation of Inductances," IEEE Trans. Power Electron., vol. 36, no. 5, pp. 5771–5783, May 2021, doi: 10.1109/TPEL.2020.3028084.

[7] H. Kim, H.-S. Jung, S.-K. Sul, and D. J. Berry, "IPMSM Magnet Temperature Estimation by d-axis Flux Linkage," in 2019 10th International Conference on Power Electronics and ECCE Asia (ICPE 2019 - ECCE Asia), Busan, Korea (South): IEEE, May 2019, pp. 2517–2522. doi: 10.23919/ICPE2019-ECCEAsia42246.2019.8796488.

[8] W. Kirchgässner, O. Wallscheid, and J. Böcker, "Thermal neural networks: Lumped-parameter thermal modeling with state-space machine learning," Engineering Applications of Artificial Intelligence, vol. 117, p. 105537, Jan. 2023, doi: 10.1016/j.engappai.2022.105537.

[9] W. Kirchgassner, O. Wallscheid, and J. Bocker, "Data-Driven Permanent Magnet Temperature Estimation in Synchronous Motors With Supervised Machine Learning: A Benchmark," IEEE Trans. Energy Convers., vol. 36, no. 3, pp. 2059–2067, Sep. 2021, doi: 10.1109/TEC.2021.3052546.

[10] W. Kirchgassner, O. Wallscheid, and J. Bocker, “Estimating Electric Motor Temperatures With Deep Residual Machine Learning,” IEEE Trans. Power Electron., vol. 36, no. 7, pp. 7480–7488, Jul. 2021, doi: 10.1109/TPEL.2020.3045596.

[11] J. Lee and J.-I. Ha, “Temperature Estimation of PMSM Using a Difference-Estimating Feedforward Neural Network,” IEEE Access, vol. 8, pp. 130855–130865, 2020, doi: 10.1109/ACCESS.2020.3009503.