IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 June 2022, accepted 12 July 2022, date of publication 18 July 2022, date of current version 21 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3191643

RESEARCH ARTICLE

Thermal Estimation of Modular Multilevel Converter Submodule Using Deep Regression on GRU and LSTM Network

YE-SEUL PARK[1], (Member, IEEE), HYE-WON CHOI[2], (Graduate Student Member, IEEE), KYO-BEUM LEE[2], (Senior Member, IEEE), AND JUNG-WON LEE[1,2], (Member, IEEE)

[1]Department of AI Convergence Network, Ajou University, Suwon-si 16499, South Korea
[2]Department of Electrical and Computer Engineering, Ajou University, Suwon-si 16499, South Korea

Corresponding authors: Jung-Won Lee (jungwony@ajou.ac.kr) and Kyo-Beum Lee (kyl@ajou.ac.kr)

This work was supported by the Korean Government (Ministry of Science and ICT, MSIT) through the National Research Foundation of Korea (NRF), the Korea Institute of Energy Technology Evaluation and Planning (KETEP), and the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea (No. 2020R1A2C1007400 and No. 20206910100160).

ABSTRACT This paper proposed a GRU/LSTM-based deep regression model for thermal estimation of modular multilevel converter submodule. The MMC is composed of many submodules with the power semiconductors such as IGBTs and MOSFETs. The switches are the main components determining the reliability of the MMCs, and the swing of junction temperature causes most switch failures in the power semiconductors. So, thermal estimation is essential to improve the reliability of the MMC systems. Thermal modeling is a regression problem of time-series data, considering various environmental conditions. The conventional models cannot reflect the complex environmental conditions due to their fixed mathematic formulas. Therefore, this paper proposes the deep regression model that can estimate the junction temperature by using the arm current of the MMC submodule. The proposed model improved the accuracy of thermal estimation by more than 7.2 times compared to the existing method. Moreover, it does not require pre-processing and takes about 4.5ms on average to process 100ms data.

INDEX TERMS Thermal estimation, modular multilevel converter (MMC), deep regression, gated recurrent unit (GRU), long short term memory (LSTM), recurrent neural network (RNN).

NOMENCLATURE

ACRONYMS	MEANING
AI	Artificial Intelligence
ANN	Artificial Neural Network
DRM	Deep Regression Model
GRU	Gated Recurrent Unit
IGBT	Insulated Gate Bipolar Transistor
LSTM	Long Short-Term Memory
MMC	Modular Multilevel Converter
RNN	Recurrent Neural Network

I. INTRODUCTION

Modular multilevel converters are used in the direct current transmission systems, energy storage systems, and motor drives for high-voltage and high-power conversion systems [1], [2]. The output voltage of MMC is decided by the number of submodules in each arm [3]. The MMC achieves high output power quality using the submodules and reduces switching losses, switching stress, and filter size compared to 2-level or 3-level inverters [4], [5]. However, failure increases when the MMC systems are exposed to harsh environments such as high voltage and medium voltage transmission systems [6]. The failure of MMC requires considerable time and cost to restoration [7], [8]. Therefore, according to the lifetime and robustness, the reliability of the MMC is still a crucial issue [9], [10].

The power conversion system consists of various components such as switches, capacitors, inductors, and gate drivers. The power semiconductor switches are the main components of power conversion systems [11]. The semiconductor device failure accounts for 21 % of converter system failure, and

The associate editor coordinating the review of this manuscript and approving it for publication was Chandan Kumar.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

the leading cause of the switches is thermal stress [12]–[14] As a result, the thermal state of the semiconductor determines the reliability and robustness of the systems [15]. The thermal stresses of the semiconductor occur in bond wire fatigue, metallization reconstruction, solder fatigue, gate oxide failure, and burnout failure [16]. Consequently, thermal estimation is essential for the reliability of power conversion systems [17], [18].

Conventional methods for thermal estimation are based on mathematical formulas: the cauer network and foster network. The cauer network is based on the actual physical properties of the switching device [19]. The cauer model has the drawback of requiring precise material parameters. On the other hand, the foster network requires the power losses of the switches by calculating the thermal time constants and resistance [20], [21]. However, the junction temperature of a submodule is affected by its power losses and that of entire systems. So, these methods have limitations to estimating thermal conditions accurately.

Thermal modeling is a regression problem of time-series data, which should consider various environmental condition. Several AI techniques have been widely used to consider various conditions [22]. [23] presented a surrogate thermal model for power electronic modules using ANNs. In [21], ANNs are used to predict the lifetime of photovoltaic systems. The established work has validated the potential of AI-aided methods supporting the analysis of power electronic systems. However, [23] has a limitation in that it created a thermal estimation model using simulation data. Thus, it is difficult to verify how robustly it operates on real data. In [24], several parameters (e.g., the switching frequency of the inverter, reference DC link, filter parameter, etc.) should be considered simultaneously while building a model. So, the complexity of a model and the amount of data processing computation increases. Reference [25] reflects the continuous characteristics of the time-series data to build AI-based model considering thermal cross-coupling (TCC) effects using a half-bridge module with two IGBTs and two diodes. But it has limitations to time and data processing because it requires data preparation process: power loss injection, thermal response measurement, and log-scale transformation.

To solve these problems, this paper proposed a thermal estimation method of MMC submodule using deep regression on GRU and LSTM network, that:

1) **is trained by gathering the measurement data** to operates on the field data robustly.

2) **only consider the arm current of the MMC submodule** so that the complexity of the model is low.

3) **uses raw data to predict** so that it does not require any data pre-processing.

Deep regression is a method of solving a regression problem using a neural network structure. This paper uses a LSTM [26] and GRU [27] which is a type of RNN and is widely used to predict time-series data [28], [29]. The proposed model does not require any pre-processing and it takes about 4.5ms on average to process 100ms data. As a result of experiment, it can estimate 1-hour data within about 164 (with GRU) and 193 (with LSTM) seconds (about 3 minutes). More encouraging experimental results showed that the accuracy of thermal estimation was improved by more than 7.2 times compared to the existing method and the prediction error of the proposed model was reduced to 1/5 of that of the existing model.

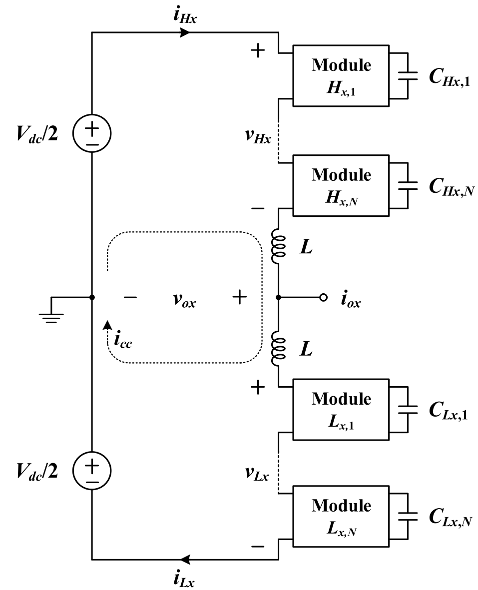

FIGURE 1. A schematic circuit of a three-phase MMC topology.

II. MATHEMATICAL MODELING

This section is discussed about the configuration of MMC submodule and mathematical model for thermal estimation. In this paper, the three-phase MMC topology is considered and there are two complementary IGBT switches and two parallel diodes on each MMC submodule. The details of the configuration are as follows:

A. MMC CONFIGURATION

Fig. 1 presents a schematic circuit of a three-phase MMC topology. Each phase of the MMC has two arms (upper and lower), and each arm is equipped with arm inductor to limit the current due to short circuit accidents. The output current i_{ox} is represented by Kirchhoff's current law as shown in equation (1), and the circulating current i_{cc} of MMC flows in a leg through both the upper and lower arms, as shown in equation (2).

$$i_{ox} = i_{Hx} - i_{Lx} = I_{ac} \sin(\omega t - \phi), \quad (1)$$

$$i_{cc} = \frac{1}{2}(i_{Hx} + i_{Lx}) = \frac{1}{3} I_{dc} + I_h \sin(2\omega t - \phi), \quad (2)$$

where the output current i_{ox} is presented by the magnitude of I_{ac} as sinusoidal function, and the circulating current i_{cc} includes the DC component I_{dc} and the harmonics. In the MMC system, the harmonic of the circulating current is

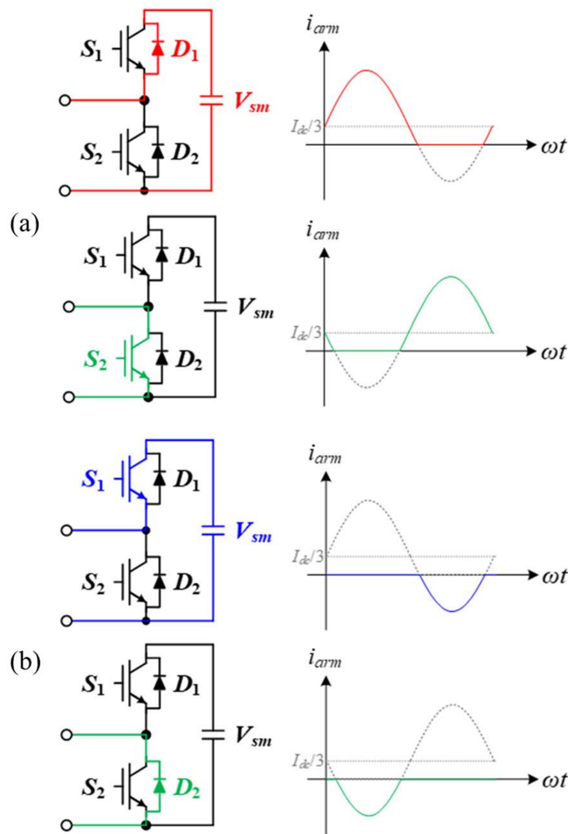

FIGURE 2. Operating modes of submodule (a) $i_{arm} > 0$, (b) $i_{arm} < 0$.

significantly small, the circulating current is expressed the DC current.

From the equations (1) and (2), the upper and lower arm current i_{Hx}, i_{Lx} are given by

$$i_{Hx} = i_{cc} + \frac{1}{2}i_o = \frac{1}{3}I_{dc} + \frac{1}{2}I_{ac}\sin(\omega t - \phi), \tag{3}$$

$$i_{Lx} = i_{cc} - \frac{1}{2}i_o = \frac{1}{3}I_{dc} - \frac{1}{2}I_{ac}\sin(\omega t - \phi). \tag{4}$$

According to Kirchhoff's voltage law, the upper and lower voltage are present as follows:

$$v_{Hx} = \frac{V_{dc}}{2} - L\frac{di_{Hx}}{dt} - v_o, \tag{5}$$

$$v_{Lx} = \frac{V_{dc}}{2} - L\frac{di_{Lx}}{dt} + v_o, \tag{6}$$

where V_{dc} represents the dc link voltage, and v_o represents the output voltage.

The upper and lower voltage are associated with the output and circulating current from the equations (3)-(6).

$$v_{Hx} - v_{Lx} = -2v_o - L\left(\frac{di_{Hx}}{dt} - \frac{di_{Lx}}{dt}\right) = -2v_o - L\frac{di_o}{dt}, \tag{7}$$

$$v_{Hx} + v_{Lx} = V_{dc} - L\left(\frac{di_{Hx}}{dt} + \frac{di_{Lx}}{dt}\right) = V_{dc} - 2L\frac{di_{cc}}{dt}. \tag{8}$$

TABLE 1. Operating modes of submodules.

Current	S_1	S_2	V_{sm}	Mode
$i_{arm} > 0$	ON	OFF	V_{dc}/N	Charging
	OFF	ON	0	Bypass
$i_{arm} < 0$	ON	OFF	V_{dc}/N	Discharging
	OFF	ON	0	Bypass

In the steady state, the variance of the output current is ideally sinusoidal, and the variance of the circulating current is negligible. The references of the upper and lower voltage are represented in equation (9) and (10).

$$v_{Hx} = \frac{1}{2}V_{dc}(1 - m\sin\omega t), \tag{9}$$

$$v_{Lx} = \frac{1}{2}V_{dc}(1 + m\sin\omega t). \tag{10}$$

The arm of each phase composes submodules, the minimum unit of the power conversion system, and the submodules generally apply the half-bridge topologies. There are two complementary IGBT switches S_1, S_2, two parallel diodes D_1, D_2, and a capacitor C_{sm} in the half-bridge modules.

The operating modes of the SMs shown in Fig. 2 are divided into three modes (charging, discharging, and bypass) depending on the arm current i_{arm} and the IGBT switching states as shown in Table 1. When the arm current flows in a positive direction, the SMs operate the charging and bypass modes. When the arm current flows in a negative direction, the SMs operate the discharging and bypass modes.

The duty function of the four power devices F_{S1}, F_{S2}, F_{D1}, and F_{D2} depend on the current flowing and the IGBT switching states, as shown in equations (11)-(14). The duty function of S_1 and D_1 are identical with the reference of the upper voltage V_{Hx}. However, S_2 and D_2 operate in a complementary to S_1 and D_1. R_{S2} and R_{D2} are equal to the reference of the lower voltage V_{Lx}.

$$F_{S1} = \frac{1}{2}(1 - m\sin\omega t) \quad i_{arm} < 0, \tag{11}$$

$$F_{D1} = \frac{1}{2}(1 - m\sin\omega t) \quad i_{arm} > 0, \tag{12}$$

$$F_{S2} = \frac{1}{2}(1 + m\sin\omega t) \quad i_{arm} > 0, \tag{13}$$

$$F_{D2} = \frac{1}{2}(1 + m\sin\omega t) \quad i_{arm} < 0. \tag{14}$$

B. MATHEMATICAL THERMAL MODELING

Mathematical thermal modeling of MMC is divided into two steps: 1) power loss calculation and 2) thermal estimation of the power devices. The energy loss E_{sw} [20] by the switching of the power devices is shown in equation (15).

$$E_{sw} = \left(E_{on} + E_{off}\right) \cdot \left\{1 + K_T \cdot \left(T_j - T_{ref}\right)\right\}, \tag{15}$$

where E_{on} and E_{off} are energy loss according to the switching state, K_T means the thermal coefficient, and T_j, T_{ref} is

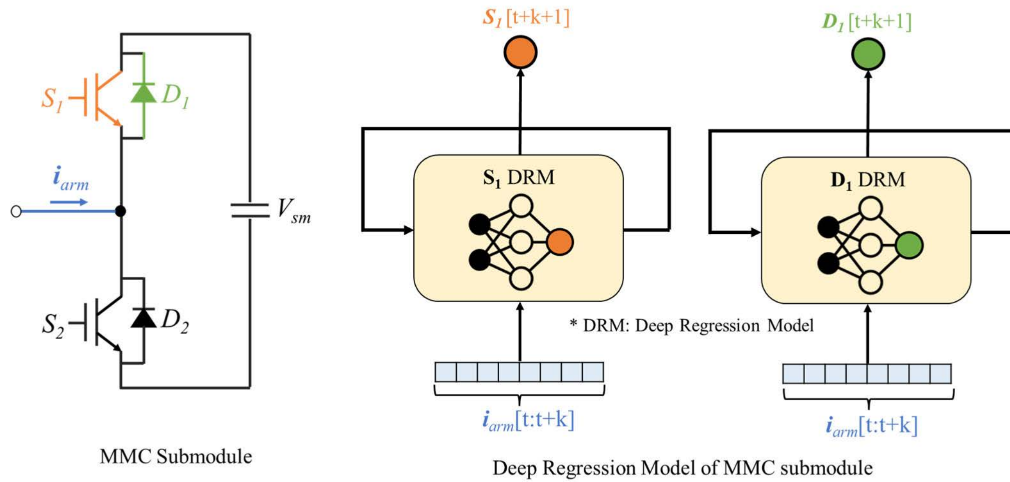

FIGURE 3. Architecture of deep regression model for thermal estimation of MMC submodule.

junction and reference temperature, respectively. The power loss by the energy loss [20] is associated with the switching frequency and the reference voltage V_{ref}, as shown equation (16).

$$P_{sw} = E_{sw} \cdot f_{sw} \cdot \left(\frac{V_{sm}}{V_{ref}}\right)^{K_V}, \tag{16}$$

where V_{sm} represents the capacitor voltage of the SMs, K_V represents the voltage coefficient. The voltage by the conduction V_{cond} is shown equation (17)

$$V_{cond} = \{V_{ref} + K_{C1} \cdot (T_j - T_{ref})\} + i_{arm} \cdot \{r_{ref} + K_{C2} \cdot (T_j - T_{ref})\}, \tag{17}$$

where r_{ref} is on-state slope resistance, K_{C1}, and K_{C2} are the temperature coefficients obtained from the datasheet. The power loss by the conduction voltage is associated with the arm current i_{arm} and the duty function F_X.

$$P_{cond} = V_{cond} \cdot i_{arm} \cdot F_X \quad (X = S_1,\ S_2,\ D_1,\ D_2). \tag{18}$$

The power loss of the power devices represents sum of the switching loss and conduction loss.

$$P_{loss} = P_{sw} + P_{cond}. \tag{19}$$

The thermal model based on a fourth-order foster network is obtained as follow [15]–[17]:

$$\Delta T_n = \sum_{k=1}^{4} \Delta T_{n-1} \cdot e^{-\frac{\Delta t}{\tau_k}} + P_{ioss} \cdot \sum_{k=1}^{4} r_k \cdot \left(1 - e^{-\frac{\Delta t}{\tau_k}}\right) \tag{20}$$

where τ_k and r_k represent time constant and thermal resistance, respectively.

III. DEEP REGRESSION ON GRU/LSTM NETWORK

This section is discussed about DRM on GRU and LSTM network. The proposed model is briefly illustrated as shown in Fig. 3. The thermal estimation is achieved by training MMC arm current ($\boldsymbol{i}_{arm}$) and temperature (S_1, D_1) data. Physically, the switch temperatures S_1 and diode D_1 change according to the i_{arm} as shown in left side of Fig. 3. Similarly, in the DRM model, the temperature of the switch(S_1) and diode(D_1) at $t + k + 1$ is predicted through the input of the arm current (i_{arm}) change from $\boldsymbol{t}$ to $t + k$. For example, assuming that the window size is 10, it means predicting the temperature at 41 seconds with arm current data between 30 and 40 seconds ($t = 30$, $k = 10$, $l = 1$). Here, using the same arm current data (i_{arm}), the temperature of the switch (S_1 DRM) and the diode (D_1 DRM) are estimated respectively. The detail is as follows:

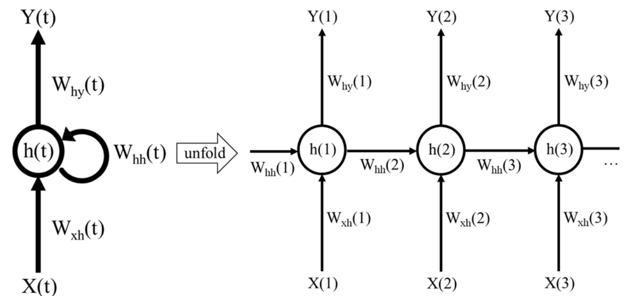

FIGURE 4. Architecture of recurrent neural network model.

A. MODEL CONSTRUCTION

Deep regression model on LSTM [26] and GRU [27] network is based on the structure of RNNs. Fig. 4 shows the basic structure of an RNN. The basic structure is on the left, which is executed repeatedly to run the model. The right side illustrates the execution flow of the model over time. The components of the model operated at an arbitrary time t are summarized as follows:

- **Input data X(t)**: X(t) is the initial data applied to the model. The type of the input data depends on the model

structure. It may be a scalar value, one or high dimensional time-series data. In this study, the current data ($\boldsymbol{i}_{arm}$) was the input data. By applying the time sliding window, X(t) was formed to a one-dimensional vector.

- **Hidden layer result h(t)**: H(t) represents the intermediate result obtained by calculating the input data, which applied at time t as the internal weight of the model.
- **Output data Y(t)**: Y(t) means the result that the model finally predicts. In this study, we designed the temperature values as scalar type of the output layer.

At this time, to calculate the results of the hidden and the output layer, the internal weight should be updated by learning. There are three types of weights in the RNN:

- **Weight in the input layer, $\mathbf{W}_{xh}$(t)**: The weight from the input layer to the hidden layer at time t.
- **Weight in the recurrent layer, $\mathbf{W}_{hh}$(t)**: The weight from the hidden layer at time t to the next time t + 1.
  - **Weight in the output layer, $\mathbf{W}_{hy}$(t)**: The weight from the hidden layer to the output layer at time t.

Training of the RNN model aims to minimize the loss of the output layer and update weights such as: $W_{xh}(t)$, $W_{hh}(t)$, and $W_{hy}(t)$. The performance of the model is also changed according to the weight connection of the hidden layer. Here, each of the three layers ($W_{xh}(t)$, $W_{hh}(t)$, and $W_{hy}(t)$) is trained with a different weight value.

B. DATA CHARACTERISTICS

The thermal estimation model is constructed by labeling data (S_1, D_1, and $\boldsymbol{i}_{arm}$). Labeling is to specifically define the input data which is the source of prediction and the output data which is the target of prediction. Here, we must concentrate on the characteristics of regression model. Building a regression model is a work to find numerical relationships between the input and the output data. If the input data to predict the same output have similar patterns, training the model will be easy. If it is difficult to predict, training the model will be difficult as well. Our model takes current data as input and temperature data as output. Let's examine the relationship between temperature and current data.

(Fig. 5. is included on the next page because of its size) In Fig. 5 (a) and Fig. 5 (b), the point A and B capture different temperature (output) values depending on whether they are rising or falling edges of the current (input) data. The difference is caused by errors in measuring temperature or slight delays in signal transmission. In contrast, the point C and D shows same temperature (output) values 42.5° of switch S_1 with the different current (input) values. When training a regression model, it is easy to predict the same output value when the input data has a uniform pattern. If the input data has different patterns with same label, it's like giving different answers to the same test question.

Fig. 5(c) clearly shows these characteristics. It shows the comparison of the input data cropped by different window sizes when predicting the same label as shown in three blue dotted lines (1∼3) in Fig. 5(b). Here, the window size means the number of the input data to be observed at once. For example, the window size of 5 means fetching the five sequential input data. In Fig. 5(c), the input data has a more uniform pattern when the window size is 100 than 5. In other words, using the input data cropped to a window size of 100 is better because it provides the insight of data trend of back and forth as to prediction basis. Therefore, we propose the data labeling method by using time sliding window, as shown in Fig. 6.

C. DATA LABELING

Fig. 6. shows two sinusoids with the same period. The gray-shading block is one period with n slots. Assuming the data set is constructed by sliding the window from the left. The first data set is built by combining the input data from three slots (1-3) and the output data from the slot 4 (green colored slot). The following second data set consists of the input data from 2-4 and the output data from 5 (yellow). The third is configure d by connecting the input data from 3-5 and the output data from 6 (orange). Here, it is important to select an appropriate window size. If the window size is too small, it operates very sensitively to small error values such as noise. If the window size is too large, a meaningful data pattern is not generated, which deteriorates the performance. Therefore, it is required to set the window size of about 0.1∼10% of a period length, compare it with the performance, and tune the model.

IV. EXPERIMENTAL RESULT

This section presents the experimental results of the proposed model. Experimental results are followed in the order of environment configuration, experimental setting, model evaluation, and discussion.

A. ENVIRONMENTAL CONFIGURATION

Fig. 7 shows a submodule of MMC experimental setup for measuring the temperature. The experimental setup consists of the following parts: a half bridge converter module (F4-50R12K24), a control board, a load reactor, dc-link window size is three and the movement size (stride) is one slot, capacitors, temperature sensors, and DC power source (TC.P.20.600.400 of REGATRON). The converter module F4-50R12K24 consists of the two half-bridge converters as SM of MMC and load converter. The current of submodule is controlled by AC 15A and DC 5A using the load converter. The silicone gel of the module is removed to accurately measure the temperature using the fiber-optic temperature sensor as shown in Fig. 8. The experiments are performed under the conditions listed in Table 2. The input voltage is applied at 300 V, and the load reactor sets to 2 mH, respectively. The switching frequency of load converter sets 10 kHz. However, the switching frequency of half-bridge converter is 1 Hz for sensing the temperature of the switches and diodes.

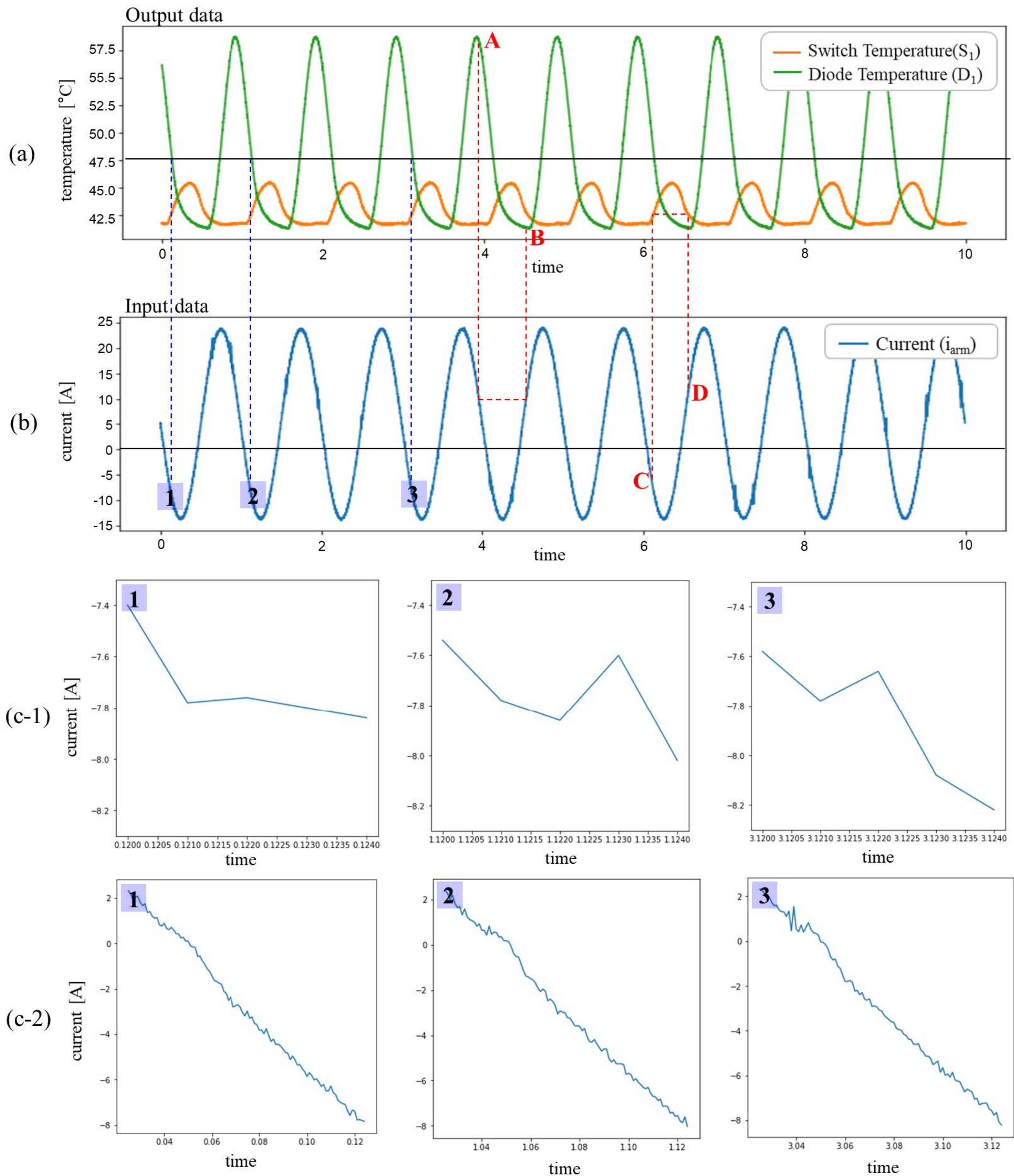

FIGURE 5. Analysis of data characteristics in multi-modular converter submodule. (a) The output data (The switch temperature: S_1, The diode temperature: D_1). (b) The input data(The arm current: i_{arm}). (c) The input data pattern according to various window size with same label (The label: The diode temperature 47.5). (c-1) The window size = 5. (c-2) The window size = 100.

B. EXPERIMENTAL SETTING

To evaluate the proposed model, the following experimental conditions were designed. The detail of the experimental setting is as follows:

- **The configuration of the model**: The model was trained with temperature data by input current of AC 15A and DC 5A. In the experiment, the mathematical model in Chapter 2 and the deep regression model in Chapter 3

FIGURE 6. Learning dataset construction method of periodic data by using time sliding window (e.g., sinusoidal signal).

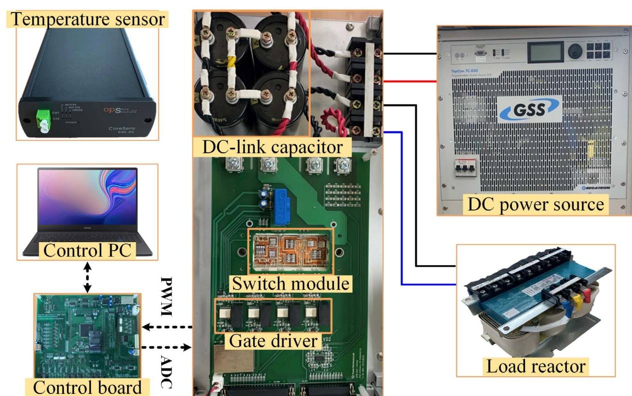

FIGURE 7. Experimental setup for measuring the temperature of MMC submodule.

were compared. There are three layers in both GRU and LSTM models. We considered the batch size and the number of neurons based on the performance metrics (R^2 score) to get the best results. The best performing case was experimentally set by adjusting the number of neurons in each layer to 32, 64, 128, and 256. We tested the 64 cases (4 × 4 × 4 cases in three layer). As a result, the number of neurons in the first layer is 64, 128 in the second layer, and 128 in the third layer. Likewise, we experimentally set the batch size, by changing 4 cases: 64, 128, 256, and 512. Consequently, the batch size is fixed to 128.

- **The ratio of the training/testing data**: It was validated with the 8:2 ratio of the training dataset and the testing dataset. With the total 400,000 frames of data, 320,000 frames were used for training and 80,000 frames were used for testing.
- **The size of the window**: Data of one period is composed of 1,000 frames. We experiment to get an optimized window size by changing the window size: from 0.1% to 10% of one cycle.

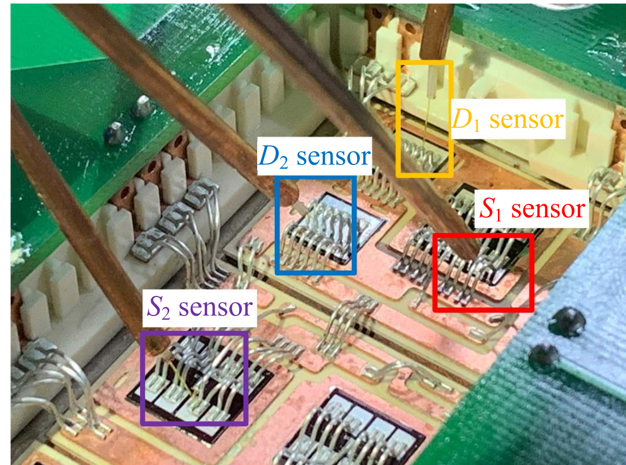

FIGURE 8. Real-time experimental setup for MMC submodule.

TABLE 2. Experimental parameters.

Parameters	S_1	Mode
DC-link voltage	300	V
Reference voltage of SM	100	V
DC-link capacitor	500	μF
Load reactor	2.5	mH
Switching frequency	10	kHz
Load current frequency	1	Hz
Sampling time of sensor	1	ms

- **The index of evaluation**: We used the R^2 score [30], MAE (Mean absolute errors) for performance evaluation. The R^2 score is widely used to measure the variance proportion of the predicted result (y) compared to the ground truth ($\hat{y}$). So, the R^2 score is suitable for evaluating the performance when learning a regression model. The MAE means the absolute loss between the predicted result (y) and ground truth ($\hat{y}$). The formula for each index is as follows (i: the index of frame, N: the total length of frame):

$$MAE = \frac{1}{N}\sum_{i}^{N}\left|y_i - \hat{y}_i\right| \tag{21}$$

C. MODEL EVALUATION

In this paper, we designed two types of models: The simulation (mathematical) model in Section 2, The GRU/LSTM-based deep regression model in Section 3. This section described the evaluation result of these model.

1) SIMULATION MODEL

Fig. 9 shows the comparative result of simulation and measurement in the MMC circuit with input current: AC 15A and DC 5A. Here, the simulation is based on Chapter 2. The blue line shows the measured data, and the orange line shows the simulation data. As shown in Fig. 9(a), there is a large

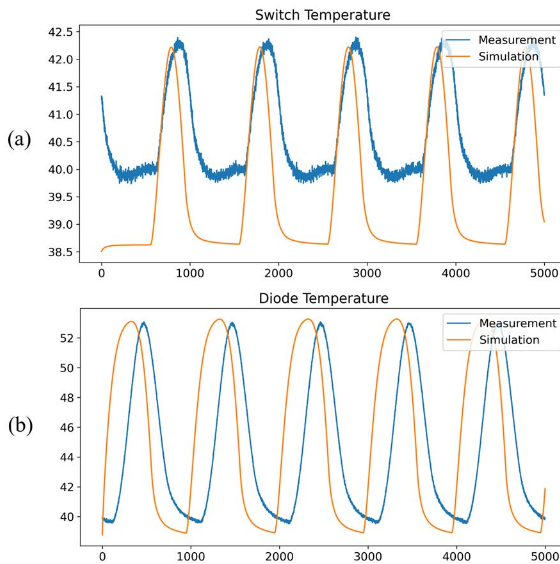

FIGURE 9. Comparison between simulation and measured data. (a) Switch (S_1). (b) Diode (D_1) temperature.

difference between the simulation and the measured data. In Fig. 9(b) shows that two graphs have some off-set along the x-axis. It shows limitations that arises because the simulation model does not reflect all conditions of the measurement environment. The MAEs of the measured and simulated data are 1.1615, and 4.9338, respectively. Now we are ready to evaluate the deep regression model for thermal estimation of MMC submodule.

2) DEEP REGRESSION MODEL (GRU/LSTM)

The DRM model is divided into the switch and diode model. For both models, three approaches (simulation, GRU and LSTM models) were compared. *(Fig. 10 ∼ Fig. 14 is included on the next page because of its size)*

- **Switch Model**: Fig. 10 and Fig. 11 show the difference between the measured temperature and the estimation result of the switch model. The simulation model had fatal errors for predicting the actual temperature regardless of the window size. GRU and LSTM models also had a problem of overfitting when the window was small. But it is stabilized making it similar to the actual data as the size of the window became larger.
- **Diode Model**: Fig. 12 and Fig. 13 show the prediction result of the diode temperature using the simulation, GRU and LSTM models. The simulation model, similar to the switch model, had significant errors in predicting the actual temperature regardless of the window size. However, unlike the switch model, this error was a time delay where the rise and fall times are faster than the measured data. GRU and LSTM model also had a problem of overfitting when the window was small.

Fig. 14 shows the estimation performance by using the R2 scores. The R2 scores of each model are compared by

TABLE 3. The analysis of the estimation error (the switch model).

Unit: (°C)	Error (min)	Error (max)	MAE
ANN [25]	-0.3	2.9	-
LSTM	-0.1091	0.7747	0.2200
GRU	-0.0425	0.7000	0.2609
Simulation	-2.6736	0.5271	1.1332

* Note: Error = Measurement – Estimation

TABLE 4. The assessment of the thermal estimation model.

	Switch			Diode	
	ANN [25]	LSTM	GRU	LSTM	GRU
Thermal Characteristics	Loss Injection	Load current Injection			
Data Pre-processing	Thermal Response	None			
Computational Time (s)	(109+α)	193	164	193	164
Average Loss (MAE)	1.6	0.22	0.26	0.42	0.42

*Conditions: Intel (R) Core (TM) i7-6700 CPU @ 3.40GHz 3.40

changing the window size from 1 to 100. The simulation model has a negative value less than 0. It means that the error between the measured value and the predicted value is larger than simply predicts with the average. In contrast, RNN than the difference between the measured value and the mean of the predicted value. In other words, it depicts an abnormal case where the model has lower performance than the model models show high performance. The performance of GRU/LSTM model changes drastically up to when the window size is 10 and it stabilizes after 10.

D. DISCUSSION

As previously analyzed, the proposed model can estimate temperature with higher performance compared to the existing mathematical model. In this section, we discuss the effect of the proposed model from three perspectives:

- **Real-time**: Real-time performance was evaluated by comparing prediction results of sensor data measured in real time. Here, we measured the actual temperature sensor data (the green line in Fig. 15) for comparison the model performance (the blue and orange line in Fig. 15). Fig. 15 shows the prediction with the current data in real-time. Fig. 15(a) shows the current data as the input for the trained model, and Fig. 15(b)-Fig. 15(c) shows the switch and diode temperature predicted by the GRU, LSTM and simulation model. In the initial part (determined by the window size) to the left of the red line, it is difficult to estimate the temperature due to the data labeling. But it is negligible time (0.1 sec). After 0.1 sec, the proposed

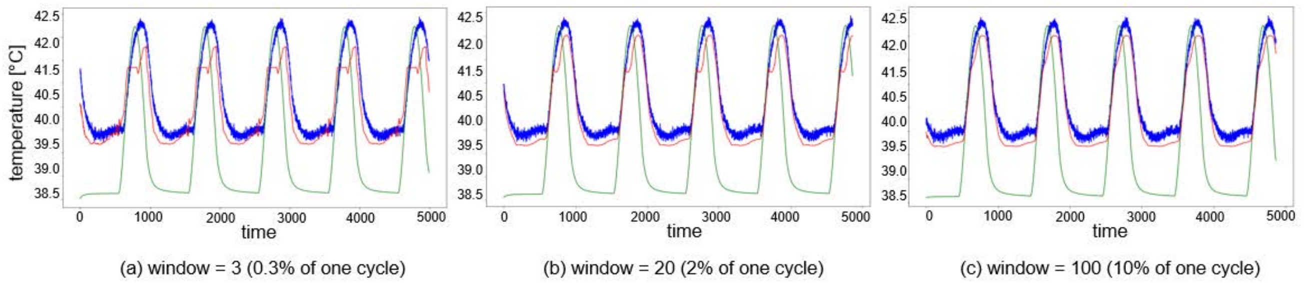

FIGURE 10. The estimation result of switch temperature (blue: measurement, green: simulation, red: GRU).

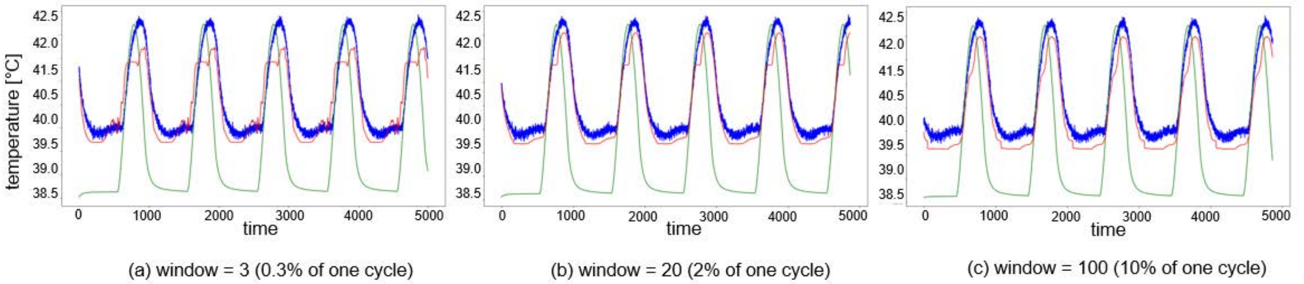

FIGURE 11. The estimation result of switch temperature (blue: measurement, green: simulation, red: LSTM).

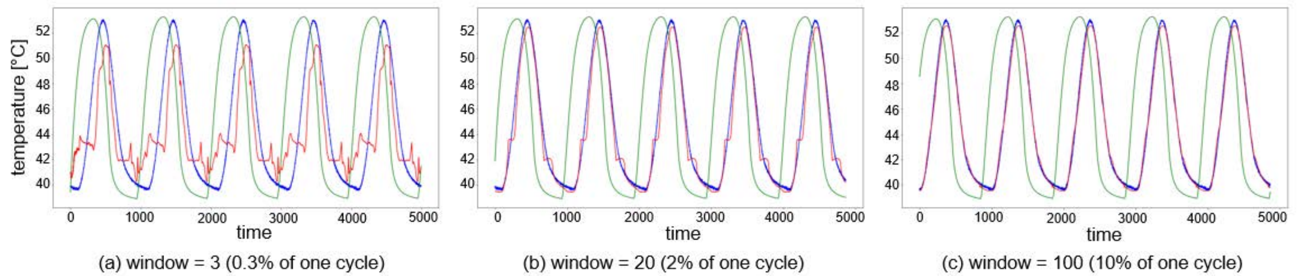

FIGURE 12. The estimation result of diode temperature (blue: measurement, green: simulation, red: GRU).

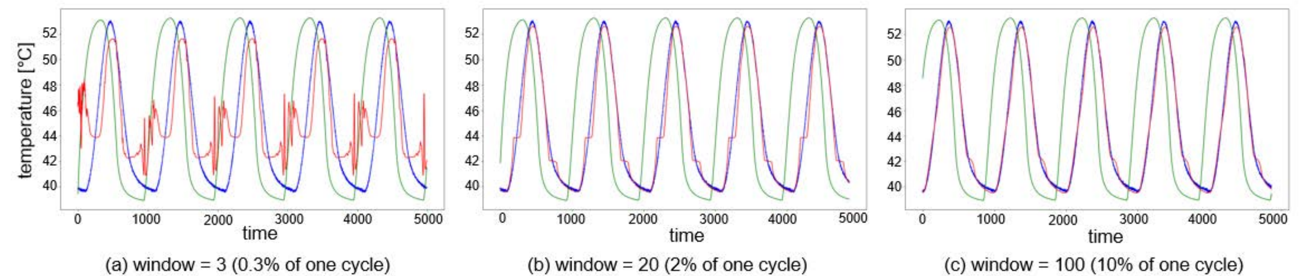

FIGURE 13. The estimation result of diode temperature (blue: measurement, green: simulation, red: LSTM).

model had high accuracy. The proposed method does not require any pre-processing task and support the real-time processing.

- **Reliability**: Table 3 shows the prediction error of the switch temperature when the window size was 100 (with the best performance). The proposed method has a small error of 0.2200 for LSTM and 0.2609 for GRU in MAE. However, the simulation model shows a difference of more than 4 times of the proposed method by MAE as 1.1342. The results were compared with the previous study [25]. It reflects the continuous characteristics of the time-series data to build

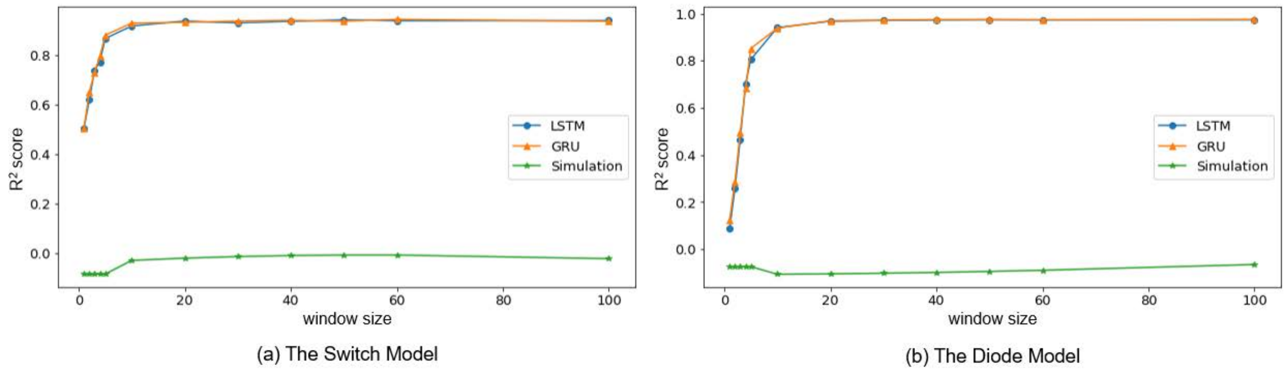

FIGURE 14. The evaluation result of the model performance by window size.

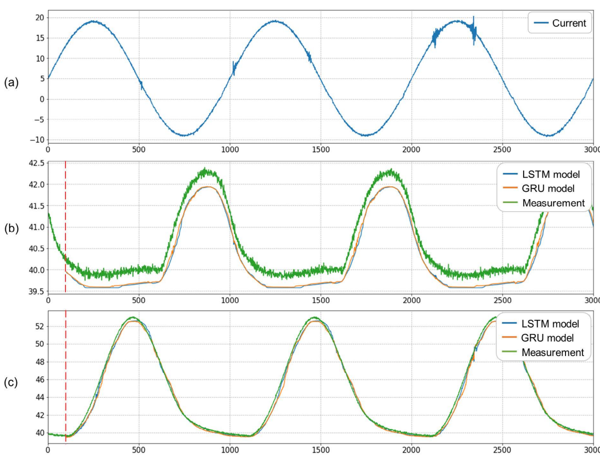

FIGURE 15. The results of thermal estimation in real-time. (a) Current, (b) switch temperature, and (c) diode temperature.

AI-based model considering thermal cross-coupling (TCC) effects, using a half-bridge module with two IGBTs and two diodes (This study used the half bridge module F4-50R12K24 and measured the temperature in the same way as in [25]). But it has limitations to time and data processing because it requires data preparation process: power loss injection, thermal response measurement, and log-scale transformation. Also, the error of the proposed model is much smaller (the accuracy of thermal estimation was improved by more than 7.2 times compared to the existing method (loss of the existing method (switch): 1.6. loss of the proposed method (switch-LSTM): 0.22). Overall, the proposed method has a stable and powerful performance at the measurement equipment level.

- **Effectiveness**: The proposed model can estimate the temperature of the switches and diodes without additional equipment or cost. Table 4 shows the time to estimate the temperature of each component with 1-hour of current data. The proposed RNN model receives real-time data by injecting an arm current and can estimate it within 193 (with LSTM) and 164 (with GRU) seconds without additional pre-processing. In addition, we entered 100ms data (the resolution of general sensors) to each model and calculated the processing time. It takes about 4.5ms on average to process 100ms data and can estimate 1-hour data within 3 minutes. Also, the average loss was found to be about 0.3, which is very low compared to the simulation model or the existing model. Compared to the previous study, the proposed model is robust in terms of computational time. For example, since temperature is feedback by thermal response, data pre-processing is required in [25]. It means that additional computational time (α) is required.

V. CONCLUSION

This paper proposes a thermal estimation method of modular multilevel converter (MMC) submodule using deep regression on a gated recurrent unit (GRU) and a long short-term memory (LSTM) network. The proposed model does not require any pre-processing. It takes about 4.5ms on average to process 100ms data and can estimate 1-hour data within 3 minutes. In addition, the accuracy of thermal estimation was improved by more than 7.2 times compared to the existing method. Consequently, the proposed model estimates the temperature of the switches at stable state without additional measurement components or costs. In the future, we plan to extend the topology of MMC submodule to the entire MMC circuit.

ACKNOWLEDGMENT

(Ye-Seul Park and Hye-Won Choi are co-first authors.)

REFERENCES

[1] K.-B. Lee and J.-S. Lee, *Reliability Improvement Technology for Power Converters*. Singapore: Springer, 2017.

[2] K.-B. Lee, *Advanced Power Electronics*. South Korea: Munundang, 2019.

[3] J. Fang, R. Blaabjerg, S. Liu, and S. M. Goetz, "A review of multilevel converters with parallel connectivity," *IEEE Trans. Power Electron.*, vol. 36, no. 11, pp. 12468–12487, Nov. 2021.

[4] Y. Ko, V. Raveendran, M. Andresen, and M. Liserre, "Thermally compensated discontinuous modulation for MVAC/LVDC building blocks of modular smart transformers," *IEEE Trans. Power Electron.*, vol. 35, no. 1, pp. 220–231, Jan. 2020.

[5] S.-M. Kim, M.-G. Jeong, J. Kim, and K.-B. Lee, "Hybrid modulation scheme for switching loss reduction in a modular multilevel high-voltage direct current converter," *IEEE Trans. Power Electron.*, vol. 34, no. 4, pp. 3178–3191, Apr. 2019.

[6] Y. Jin, Q. Xiao, H. Jia, Y. Mu, Y. Ji, T. Dragieevic, R. Teodorescu, and F. Blaabjerg, "A novel sliding-discrete-control-set modulated model predictive control for modular multilevel converter," *IEEE Access*, vol. 9, pp. 10316–10327, 2021.

[7] S.-R. Jo, S.-M. Kim, S. Cho, and K.-B. Lee, "Development of a hardware simulator for reliable design of modular multilevel converters based on junction-temperature of IGBT modules," *Electronics*, vol. 8, no. 1127, pp. 1–15, Oct. 2019.

[8] J. Sheng, H. Yang, C. Li, M. Chen, W. Li, X. He, and X. Gu, "Active thermal control for hybrid modular multilevel converter under overmodulation operation," *IEEE Trans. Power Electron.*, vol. 35, no. 4, pp. 4242–4256, Apr. 2020.

[9] S.-M. Kim, M.-G. Jeong, J. Kim, and K.-B. Lee, "Hybrid modulation scheme for switching loss reduction in a modular multilevel high-voltage direct current converter," *IEEE Trans. Power Electron.*, vol. 34, no. 4, pp. 3178–3191, Apr. 2019.

[10] J. Zhao, F. Deng, W. Hu, Y. Du, and S. Abulanwar, "Thermal optimization strategy based on second-order harmonic circulating current injection for MMCs," *IEEE Access*, vol. 9, pp. 80183–80196, 2021.

[11] Y.-J. Kim, S.-H. Kim, S.-M. Kim, and K.-B. Lee, "Open fault diagnosis and tolerance control for grid-connected hybrid active neutral-point-clamped inverters with optimized carrier-based pulse width modulation," *IEEE Access*, vol. 8, pp. 145542–145551, 2020.

[12] T. Chen, Y. Pan, and Z. Xiong, "Fault diagnosis scheme for single and simultaneous open-circuit faults of voltage-source inverters on the basis of fault online simulation," *J. Power Electron.*, vol. 21, no. 2, pp. 384–395, Jan. 2021.

[13] A. Abuelnaga, M. Narimani, and A. S. Bahman, "Power electronic converter reliability and prognosis review focusing on power switch module failures," *J. Power Electron.*, vol. 21, no. 6, pp. 865–880, Mar. 2021.

[14] N. B. Kadandani, M. Dahidah, S. Ethni, and M. Muhammad, "Lifetime and reliability improvements in modular multilevel converters using controlled circulating current," *J. Power Electron.*, vol. 21, no. 10, pp. 1611–1620, Aug. 2021.

[15] C. Guo, S. Zhang, L. Wei, H. Li, S. Wang, and K. Zhu, "Junction temperature measurement method for IGBTs using turn-on Miller plateau duration," *J. Power Electron.*, vol. 21, no. 9, pp. 1374–1382, Jul. 2021.

[16] S. Yang, D. Xiang, A. Bryant, P. Mawby, L. Ran, and P. Tavner, "Condition monitoring for device reliability in power electronic converters: A review," *IEEE Trans. Power Electron.*, vol. 25, no. 11, pp. 2734–2752, Nov. 2010.

[17] K. Ma, M. Liserre, F. Blaabjerg, and T. Kerekes, "Thermal loading and lifetime estimation for power device considering mission profiles in wind power converter," *IEEE Trans. Power Electron.*, vol. 30, no. 2, pp. 590–602, Feb. 2015.

[18] F. Hahn, M. Andresen, G. Buticchi, and M. Liserre, "Thermal analysis and balancing for modular multilevel converters in HVDC applications," *IEEE Trans. Power Electron.*, vol. 33, no. 3, pp. 1985–1996, Mar. 2018.

[19] K. Ma, N. He, M. Liserre, and F. Blaabjerg, "Frequency-domain thermal and characterization of power semiconductor devices," *IEEE Trans. Power Electron.*, vol. 31, no. 10, pp. 7183–7193, Oct. 2016.

[20] Y. Zhang, H. Wang, Z. Wang, Y. Yang, and F. Blaabjerg, "Simplified thermal modeling for IGBT modules with periodic power loss profiles in modular multilevel converters," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2323–2332, Mar. 2019.

[21] N.-C. Sintamarean, F. Blaabjerg, H. Wang, F. Iannuzzo, and P. Rimmen, "Reliability oriented design tool for the new generation of grid connected PV-inverters," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2634–2635, May 2015.

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

[23] Z. Xu, Y. Gao, X. Wang, X. Tao, and Q. Xu, "Surrogate thermal model for power electronic modules using artificial neural network," in *Proc. IECON 45th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2019, pp. 3160–3165.

[24] T. Dragičević, P. Wheeler, and F. Blaabjerg, "Artificial intelligence aided automated design for reliability of power electronic systems," *IEEE Trans. Power Electron.*, vol. 34, no. 8, pp. 7161–7171, Aug. 2019.

[25] Y. Zhang, Z. Wang, H. Wang, and F. Blaabjerg, "Artificial intelligence-aided thermal model considering cross-coupling effects," *IEEE Trans. Power Electron.*, vol. 35, no. 10, pp. 9998–10002, Oct. 2020.

[26] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, 2000.

[27] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*.

[28] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 240–255, Mar. 1994.

[29] G. Dorffner, "Neural networks for time series processing," *Neural Netw. World*, vol. 6, no. 4, pp. 447–468, 1996.

[30] G. J. Hahn, "The coefficient of determination exposed," *Chemtech*, vol. 3, no. 10, pp. 609–612, 1973.

YE-SEUL PARK (Member, IEEE) received the B.S. and M.S. degrees in electrical and computer engineering and the Ph.D. degree in AI convergence networks from Ajou University, Suwon-si, South Korea, in 2015, 2017, and 2022, respectively. Her research interests include knowledge engineering, data analysis, artificial intelligence, and prognosis health management (PHM).

HYE-WON CHOI (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in electrical and computer engineering from Ajou University, Suwon-si, South Korea, in 2019 and 2021, respectively, where she is currently pursuing the Ph.D. degree in electrical and computer engineering. Her current research interests include DC/DC converters and electric vehicle applications and reliability.

KYO-BEUM LEE (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical and electronic engineering from Ajou University, Suwon-si, South Korea, in 1997 and 1999, respectively, and the Ph.D. degree in electrical engineering from Korea University, Seoul, South Korea, in 2003. From 2003 to 2006, he was with the Institute of Energy Technology, Aalborg University, Aalborg, Denmark. From 2006 to 2007, he was with the Division of Electronics and Information Engineering, Jeonbuk National University, Jeonju-si, South Korea. In 2007, he joined the Department of Electrical and Computer Engineering, Ajou University. His research interests include electric machine drives, renewable power generations, and electric vehicle applications. He is the Editor-in-Chief of the *Journal of Power Electronics* and an Associate Editor of the IEEE TRANSACTIONS ON POWER ELECTRONICS.

JUNG-WON LEE (Member, IEEE) received the B.S. and M.S. degrees in computer science and the Ph.D. degree in computer engineering from Ewha Womans University, Seoul, South Korea, in 1993, 1995, and 2003, respectively. She worked as a Research Engineer with LG Electronics, from 1995 to 1997; and completed an internship at the IBM Almaden Research Center, Data Mining Group, San Jose, CA, USA, in 2000. She was a Research Professor and a full-time Lecturer with Ewha Womans University, from 2003 to 2006. In 2006, she joined the School of Electrical and Computer Engineering and AI Convergence Network, Ajou University, Suwon-si, South Korea. Her research interests include context awareness, big data analysis and predictive maintenance, collaborative robots, and intelligent embedded software.