IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 27, 2021, accepted November 8, 2021, date of publication November 11, 2021,
date of current version November 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3127542

Bayesian Deep Neural Network to Compensate for Current Transformer Saturation

SOPHEAP KEY[1], SANG-HEE KANG[1], (Member, IEEE), NAM-HO LEE[2], (Member, IEEE), AND SOON-RYUL NAM[1], (Member, IEEE)

[1]Department of Electrical Engineering, Myongji University, Yongin 17058, South Korea
[2]Korea Electric Power Research Institute, Daejeon 34056, South Korea

Corresponding author: Soon-Ryul Nam (ptsouth@mju.ac.kr)

This work was supported in part by Korea Electric Power Corporation (KEPCO) under Grant R17XA05-2, and in part by the Korea Research Foundation funding from the Government (Ministry of Education), in 2019, under Grant NRF-2019R1F1A1059619.

ABSTRACT Current transformer saturation has a negative effect on the operation of IEDs, resulting in their malfunction. Here, we present a technique to compensate for saturated waveforms using Bayesian Deep Neural Network (BDNN) comprising Deep Neural Network (DNN) and Bayesian optimization (BO). DNN, that utilizes stacked denoising autoencoder (SDAE) and Backpropagation (BP), is employed to optimize deep learning structure. Unlike the conventional neural network, which is a shallow network or random-initialize weights, the SDAE calculates optimal weights for each hidden layer and BP uses them to fine-tune which yields results with high performance for CT saturation compensation. To improve the empirical search of training hyperparameters, Bayesian optimization is adopted to decide training-related vectors such as batch size, learning rate, and number of neurons. Finally, the performance of the proposed approach was evaluated on an overhead transmission line which is imported from PSCAD/EMTDC with the different scenarios of fault inception angle, remnant flux, and voltage system. Therefore, numerical cases of saturation were comprehensively evaluated to demonstrate the performance of the proposed algorithm. A comparative analysis was shown to demonstrate that the proposed BDNN is superior to artificial neural network (ANN), and least square error (LES) technique.

INDEX TERMS Current transformer, saturation, deep learning, stacked denoising auto-encoders, Bayesian.

NOMENCLATURE

The main parameters used throughout this paper are summarized here.

i_P	Primary current.
i_S	Secondary current.
i_M	Magnetizing current.
L_M	Magnetizing inductance.
R	Secondary resistance (burden).
i_{AC}	Fault signal.
i_{DC}	DC component signal.
A_k	Amplitude of the k^{th} harmonic component.
φ_k	Phase angle of the k^{th} harmonic component.
M	Highest order of harmonic in the signal.
A_0	Magnitude of DC component.
τ	Time constant of the system.
$\lambda(t)$	Flux induced in the CT core.
N	Number of turns in secondary winding.
S_{CT}	Saturation slope.
A_{signal}	Magnitude of signal.
A_{noise}	Magnitude of random noise.
W_{en}	Weight of encoding layer.
b_{en}	Bias of the encoding layer.
f_{en}	Activation function of the encoding layer.
W_{de}	Weight of decoding layer.
b_{de}	Bias of the decoding layer.
f_{de}	Activation function of the decoding layer.
L	Cost function of the autoencoder.
X'	Normalized value of the input datasets.
x_{max}	Maximum values of input dataset.
x_{min}	Minimum values of input dataset.
m	Window dimension.
s	Moving step.
$f(x)$	Objective function for Bayesian optimization.

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

I. INTRODUCTION

Current transformers (CTs) are used to convert a high-voltage primary current to a relatively lower-voltage secondary

current that can be easily read and isolate the circuit from fault conditions. Reference [1] presents a detailed discussion of CT functionality and provides many useful examples. Fault current, system complexity, and size may rise due to the expansion of the interconnected power systems. Therefore, CT saturation is inevitable because of the large proportion of the fault current. CT saturation is a significant problem for power systems as it provides the incorrect current magnitude to CT. High fault currents, X/R ratio, and remnant flux are commonly contributing factors to CT saturation which in turn may lead to malfunctions in the protective relay [10]. Therefore, a robust CT saturation compensation scheme is required to clear it rapidly and reliability in order to avoid the system disturbance and equipment from damage.

This problem has been addressed in many approaches in recent years [2]–[14]; one such algorithm is a conventional neural network. The undistorted waveform was reproduced by Artificial neural network (ANN) from the saturated behavior; however, this approach was merely focused on saturation compensation without considering the remnant flux [2]. Another ANN-based method in [3] for CT saturation compensation was presented which included the remnant flux in the core and other influence parameters for CT saturation. However, the training structure in [2], [3] was empirically decided that is very time-consuming. A modified version of Adaptive Neuro Fuzzy Inference System (ANFIS) with least-square method and gradient descent, was applied on CT saturation to compensate saturated fraction [4]. However, the drawback of this algorithm is computation burden when there is large input dimension. In [5], CT saturation was successfully removed by estimating the magnetizing current from the negative value of the second difference function and adding it to the measured secondary current. However, this approach is limited by its reliance on the magnetization curve of a specific CT. A hybrid algorithm comprising the partial sum (PS) and multistage least-square (MLS) methods was proposed to address for DC offset and CT saturation problems, and was able to successfully solve these issues with a quicker response time [6]. Alternatively, two-level compensation filters can be used to compensate for saturation effects and reproduce an undistorted waveform from fault current and an inrush current with a low error [7]. The integration CT saturation detection and compensation was proposed in [8] based on sample-based extraction from the identified unsaturated samples using Kalman filter and simply reconstructed with wave shape properties and fault current characteristics. Another compensation approach utilized a least-error square (LES) filter to estimate the phasor parameters of the CT secondary current, and CT burden [9]. In [10], the unsaturated portions are extracted from distorted secondary current and then utilize the least square curve fitting method to estimate parameters for compensating CT saturation. However, due linearization by Taylor series expansion, it produces some error when the time constant is small. The application of the wavelet transforms for compensating the saturated signal was proposed in [11]. Extended Kalman filter was used to detect and compensate CT saturation by using the current sample points during the unsaturated regions to estimate an appropriate model [12]. With this estimated model, it reconstructed the signal from unsaturated behavior. The LES filter and lookup table (LUT) was jointly applied to solve DC offset and CT saturation problem which renders superb results in the presence of noise and harmonic [13]. The author of [14] proposed a new algorithm to detect and compensate for CT saturation based on derivative of secondary current and Newton's backward difference. However, this algorithm is really relied on low-pass filter characteristic to detect saturation.

Recent advances in deep learning has been widely used in power systems and energy applications because of its robustness, speed, and powerful learning capacity. As proposed in Reference [18], DC offset was generated with the random noise and several harmonics. Different scenarios such as voltage level, and inaccurate time constant, were considered to display the performance of the proposed method. Thus, results outperform the conventional filters in terms of speed and accuracy. The author of [19] used the unsupervised learning feature extraction to classify severity of CT saturation and it yielded superb results compared to conventional regression methods. Reference [20] highlighted a combining model of unsupervised feature learning and convolution neural network (CNN) to detect and classify faults based on the three-phase voltage and current signals. The discrimination between inrush current and magnetizing current in power transformer using the CNN, fast GRNN, and CLGNN is proposed in [21]–[24]. Although the method in [24] is useful to compensate for CT saturation, it requires detecting the start and end of CT saturation prior to compensation.

To the best of our knowledge, SDAE has rarely been applied on energy and power systems. One related work is a SDAE model for CT saturation compensation with the empirical selection of the training hyperparameters studied in [17]. This paper developed DBNN-based method which uses deep neural network (DNN) and Bayesian optimization (BO) to expand the capability of compensating for CT saturation. On the basis, DNN's structure is constructed based on SDAE in an unsupervised manner to extract the feature of the input during the pre-training and initialize the appropriate weights for the structures. The reconstruction of unsaturated waveform is obtained by performing backpropagation on trained model from SDAE in a supervised manner with a given label to compensate for CT saturation. Moreover, the training hyperparameters are optimized via BO to reduce difficulty for the empirical selection of training hyperparameters [29]. Lastly, we validate BDNN on simulated data from PSCAD/EMTDC with real power system parameters on a typical Korean transmission line and compared with conventional methods (ANN, and LES), study the impact of BDNN under various CT saturation conditions, and compare with other intelligent methods and conventional methods from Ref. [10]. To summarize, this paper makes the following contributions:

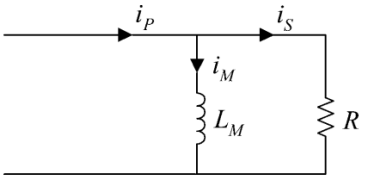

FIGURE 1. A Simplified equivalent circuit of CT.

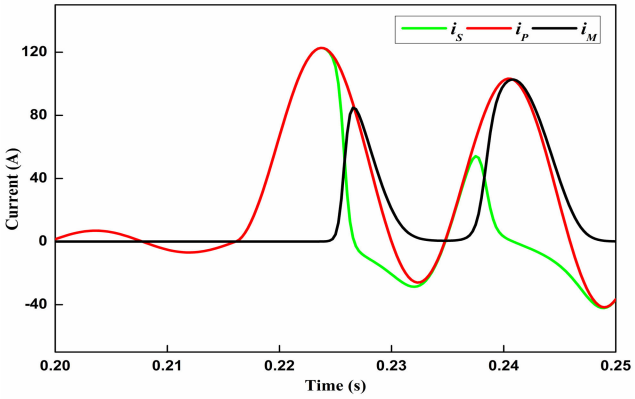

FIGURE 2. An example of CT saturation on 154kV system (0° fault angle and remnant flux of 60%).

- Propose a BDNN framework for CT saturation compensation for the first time using SDAE to extract feature that can accurately correct the distorted waveform and handle noise problem without using low-pass filter.
- Adopt Bayesian optimization to search for training hyperparameters.
- Compensate for CT saturation without the detection of CT saturation.
- Provide a consistency and high accuracy on different voltage in power systems.

The remainder of this paper is organized as follows. Section II reviews the CT saturation literature, Section III describes the proposed compensation method which comprises of adopting pre-training and fine-tuning, and data preparation. In section IV, the simulation results and performance evaluation were presented through many cases of the problem after training and the comparative study with other methods is also displayed. Section V presents concluding remarks.

II. PROBLEM STATEMENT

This section first highlights the CT saturation formulation and then introduces CT saturation datasets generation for training procedure.

A. CT SATURATION FORMULATION

A simplified equivalent circuit of a CT with a pure resistive burden is given in Fig. 1. Fig. 2 shows an example of CT saturation with primary, secondary and magnetizing current. Under the normal operation, the magnetizing current is negligible because the exciting voltage is less than knee-point. When the exciting voltage exceeds the knee-point, CT begins to experience saturation which negatively causes the magnetizing current to increase according to the B-H hysteresis curve of the CT core. Therefore, primary and secondary current are no longer proportionally equal due to the induced magnetizing current. The saturation severity is determined by the excitation current magnitude, and the saturation duration is a function of the X/R ratio. In this paper, we simplified the equations for CT saturation from [6]. The fault inception angle determines whether the saturation is positive or negative. The primary current is the sum of i_{AC} and i_{DC}, and it can be expressed as follows:

$$\begin{aligned} i_P(t) &= i_{AC}(t) + i_{DC}(t) \\ &= \sum_{k=1}^{M} A_k sin(k\omega t + j_k) + A_0 e^{-\frac{t}{\tau}} \end{aligned} \tag{1}$$

$\lambda(t)$ is proportional to i_S and it can be expressed as follows.

$$\lambda(t) - \lambda(t-1) = \frac{1}{N} \int_{t_0}^{t} R i_S(t) dt \tag{2}$$

By using the magnetizing curve, (2) can be simply rewritten as:

$$S_{CT}(i_M(t) - i_M(t_0)) = \int_{t_0}^{t} R i_S(t) dt \tag{3}$$

However, the value $i_M(t_0)$ is approximately zero before saturation begins as illustrated in Fig. 2 before fault occurs and $i_M(t)$ at the first saturation instant can be subsequently given by:

$$i_M(t) = \frac{1}{S_{CT}} \int_{t_0}^{t} R i_S(t) dt \tag{4}$$

i_S is the difference between i_P and i_M and it is given in (5).

$$i_S(t) = i_P(t) - i_M(t) \tag{5}$$

B. CT SATURATION DATASET FOR TRAINING PROCEDURE

BDNN approach is a data-dependent algorithm which requires many training datasets to obtain best results. Saturated and unsaturated data are generated by well-documented sheet from the IEEE Power System Relaying Committee (PSRC) [26]. Core-induced flux and magnetizing current are thoroughly computed in the excel sheet. Different saturation cases were obtained by modifying some key parameters (primary fault current, load resistance, X/R ratio, DC offset units, and remnant flux). The saturation becomes severe at the fault inception angle of 0°, and remnant flux of 60%, and lasts longer when X/R is 18. The training parameters in Table 1 are used to generate training input samples and it includes many types of saturation ranged from light to severe saturation. The training data for this study is 15680 saturation cases which accumulates approximately 11038720 datasets. The saturated current with random noise is used as an input, and the unsaturated current is used as the network label. Moreover, A_{noise} is added to the input A_{signal}

TABLE 1. Training data parameters.

Parameters	Values
Fault inception angle (^{0})	0,30,45,60,90
Remnant flux (%)	0,10,20,30,40,50,60
X/R	6,8,10,12,14,16,18
Burden (ohm)	2,4
Noise (dB)	20,30
2nd harmonic (%)	0,20
3rd harmonic (%)	0,13
4th harmonic (%)	0,6
5th harmonic (%)	0,2

and its magnitude is generated by modifying a signal-to-noise ratio (SNR), which can be expressed in Equation (6). Harmonic from 2nd to 5th order was also considered in the training dataset and its magnitudes are shown in Table 1.

$$SNR_{dB} = 10log(A_{signal}/A_{noise}) \tag{6}$$

III. CT SATURATION COMPENSATION APPROACH

CT saturation compensation aims to reconstruct undistorted waveform under various scenarios of power system such as power system voltage, fault inception angles, and fault types. The proposed framework (BDNN) is based on BO and DNN. SDAE extracts features in hidden layers and use it to build a deeper DNN structure. To improve the compensation accuracy, BO is employed prior to SDAE to globally optimize the training hyperparameters such as learning rate, batch size, and number of neurons.

A. THE FRAMEWORK OF DENOISING AUTOENCODER

A denoising autoencoder is an unsupervised ANN which uses nonlinear feature extraction to reconstruct clean inputs from noisy inputs, and efficiently compress and decodes data [27]. In a simple autoencoder, input $x \in R_n$ $(x_0, x_1, x_2, \ldots, x_n)$ includes in the training dataset. The input is then encoded to low dimension and it is restored to its original structure in the decoding part. The training uses the BP to minimize the reconstruction error between input and output until desired epoch is reached. After the training convergence, the encoder model and extracted features $(f_1, f_2, \ldots, f_n)$ are saved which is used to train other autoencoders. Given the input x, the output vector of the autoencoder $\hat{x}$ is mathematically expressed as follows.

$$\hat{x} = f_{de}(W_{de}f_{en}(W_{en}(x) + b_{en}) + b_{de}) \tag{7}$$

where f_{en} is Leaky ReLu activation function for encoding layer. Linear function is the decoding layer activation function f_{de} for regression task. By incorporating Adam optimizer, the training can be proceeded. Root mean square error (RMSE) is used to minimize the error between the input and output.

$$L(\hat{x}, x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N-1}(\hat{x}_i - x_i)^2} \tag{8}$$

To yield the result from autoencoder, we minimize cost function L by iteratively updating weights and bias values by using backpropagation [25]. The result of autoencoder is obtained when L is converged to a certain iteration.

B. ESTABLISHMENT OF PROPOSED BDNN

Traditionally, achieving deeper neural network structure was a very challenging task because the computation error in backpropagation process tends to dramatically increase when the structure becomes very complex (vanishing gradient problem). To strengthen the conventional neural network, authors of [15], [16] presented a new way of network training by adopting an unsupervised pre-training instead of manually selection the neural network parameters. The algorithm employed a layer-by-layer unsupervised learning based on the deep belief network (DBN) in which an unsupervised greedy layer-wise training was proposed to provide an optimization for the deeper structures. All layer parameters that are initialized during pre-training are tuned in the final stage in order to significantly achieve great results. Instead of using DBN, SDAE is adopted to construct deeper networks and it yields improved results [27]. Unlike conventional deeper networks, SDAE is used to reduce the complexity of error estimation by forming hidden layer one at a time. In a deeper structure, training efficiency is affected by initial weights; therefore, this problem can be alleviated by adopting SDAE. The main idea is to train one layer at a time by minimizing the reconstruction error. The feature of the *i-th* hidden layer is used as input for the *(i+1)-th* hidden layer.

The first autoencoder is trained in a bottleneck fashion with initial weights and bias (w_1, b_1). The distorted waveform x with the random noise is then transformed into a low dimension through encoding function and restored back to its original dimension in the decoding layer. The optimal is obtained when the error function (L) reaches the minimum and it is shown in (7). After converging to the minimum, the hidden layer, which is so-called abstract features, is stored and used as the input for the second autoencoder. After removing the decoding layer $\hat{x}$ in the first autoencoder, a new hidden layer h_2 and output $\hat{h}_1$ are stacked onto the first autoencoder as shown in Fig. 3. Using a similar process, many autoencoders were successively stacked together to form a deeper network structure. This process of using stacked autoencoders is commonly referred to as pre-training because it resembles as Restricted Boltzmann Machine (RBM). Lastly, prior layer is trained with the given label (undistorted waveform) at the output layer to reconstruct unsaturated signal. All optimal SDAE weights and bias (w_i, b_i, and $i = 1, 2, \ldots, n$) which are obtained during the pre-training were fine-tuned by backpropagation algorithm to achieve significant results in the fine-tuning. Fig. 3 presents the proposed methodology in depth.

C. PRE-PROCESS OF INPUT DATASETS

Table 1 concretely generates one-dimension saturated waveform. Due to different variation of input magnitude, BDNN

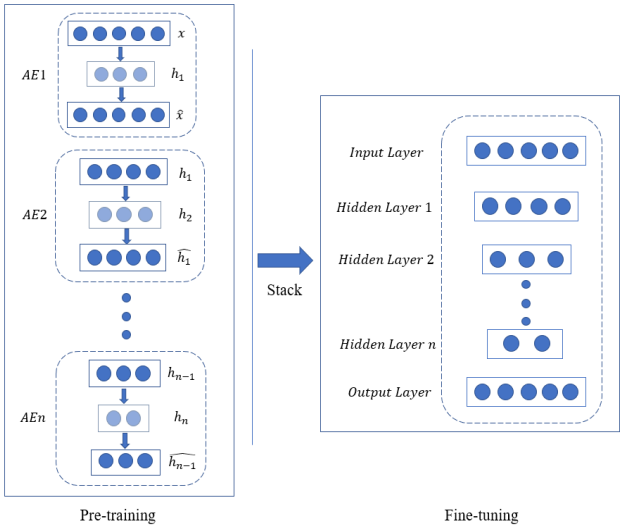

FIGURE 3. The overall structure of the proposed framework using stacked autoencoders training.

might produce large error and renders output inconsistently; therefore, it requires to be normalized to the similar magnitude to reduce the influence of this variation. As a result, it accelerates DBNN training and improve generalization of neural network [31]. The normalization formula is given as follows:

$$X' = \frac{2 \times (x_i - x_{\min})}{x_{\max} - x_{\min}} - 1 \tag{9}$$

where x_i is the i_{th} sample of the input data. After normalization is conducted, a moving-window algorithm is applied to input normalized datasets to form DBNN input training matrix. Selecting m and s are vital to reconstruct unsaturated waveform. In this study, we set m to 64 that is equal to number of samples per cycle with s of 1. Considering an input $x = \{x_1, x_2, \ldots, x_i\}$, where i is the input index. By using x, input training matrix can be formed as follows where N is the number of signal index.

$$X' = \begin{bmatrix} x_1 & x_2 & \cdots & x_{m-1} & x_m \\ x_2 & x_3 & \cdots & x_m & x_{m+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{N-m-s+1} & x_{N-m-s+2} & \cdots & x_{N-s-1} & x_{N-s} \end{bmatrix} \tag{10}$$

D. DEEP LEARNING HYPERPARAMETERS TUNING

Searching for the optimal training hyperparameters is normally performed by randomly choosing a set of hyperparameters which is very time-consuming to attain these hyperparameters. In recent years, grid search and random search are the most common optimization for hyperparameter tuning in machine learning and deep learning. However, as the dimension of tuned hyperparameters increases, the search of optimal hyperparameters in the grid search increases exponentially. Random search is proposed to deal with the problem occurred in the grid search where the random combinations of the hyperparameters are used to find the best solution for the model. For a huge dataset, it is time-consuming to achieve the optimal hyperparameters in the random search. Therefore, Bayesian optimization (BO) comes in as a tool to efficiently tune machine learning and deep learning hyperparameters [29] that chooses the hyperparameters giving in more optimal solution. Thus, BO is chosen as a hyperparameter tuning for SDAEs during the pre-training to solve this searching difficulty.

Bayesian optimization (BO) is an effective global optimization of black-box functions which is based on a probabilistic model (gaussian process) to measure the objective function in the search space. Bayesian surrogate model helps to represent the underlying objective function of the problem and acquisition model selects the next evaluation point based on prior knowledge. Our goal with respect to BO was to identify the best combination of hyperparameters x_t for an objective function $f(x)$, the combination that maximizes the output of a given search space X. The search space is designed to be a 3-dimension vector (learning rate, batch size, and number of neurons). Then, a surrogate model is built for optimization process to estimate a set of hyperparameters for SDAE with an initial hyperparameter set by distributing the gaussian process that matches the similarity of hyperparameters in the search space and use it as a prior knowledge for the optimization process. After that, we choose Expected Improvement (EI) for the acquisition function which optimizes the locations in the search space by using the prior knowledge to generate the next samples for evaluation. EI chooses the next point x_n in the search space X to evaluate that yields the smallest error. After several iteration, the best training hyperparameters can be obtained from BO.

$$x_t = arg\, max_{x \in X} EI(x) \tag{11}$$

$$\text{with } EI(x) = E(max(f(x) - f^*, 0)) \tag{12}$$

where, f^* is the maximum value that $f(x)$ has experienced during the optimization process. The flowchart of the proposed BDNN is given in Fig. 4.

IV. PERFORMANCE EVALUATION

A. DATA GENERATION FOR TESTING PROCEDURE

To evaluate BDNN efficiency, a typical three-phase overhead transmission line as shown in Fig. 5 is modelled in PSCAD/EMTDC which generates cases for BDNN testing. A CT model in [30] is utilized to generate saturation data for training in PSCAD/EMTDC with a ratio of 2000:5 (C400, $R_2 = 0.61\Omega$) and a resistive burden of 3.42Ω. Phase A current at the relay point is collected to evaluate with various fault inception angle and remnant flux. The sampling frequency is set to 3840Hz or 64 samples cycle in a 60Hz system. Then, the imported signal will pass through the moving-window technique in (10) to create the testing datasets for BDNN. When pre-processing is correctly configured, the training process is conducted on a graphics

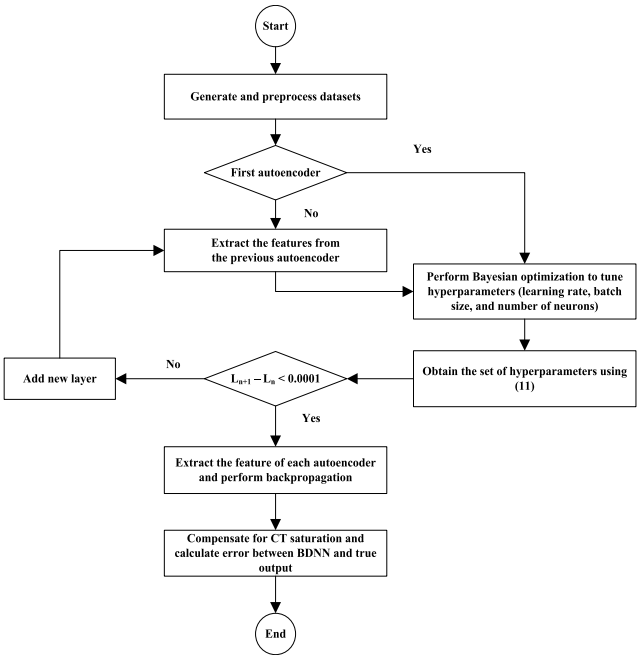

FIGURE 4. Flowchart of the proposed BDNN.

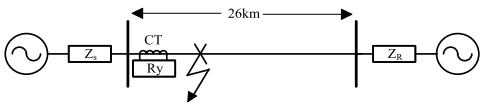

FIGURE 5. Single line diagram of 154kV and 345kV overhead transmission system.

TABLE 2. Overhead transmission parameters.

Sequence	Parameters	Value		Unit
		154kV	345kV	
Positive & Negative	R_1, R_2	0.0419	0.0169	Ω/km
	L_1, L_2	0.8921	0.6617	mH/km
	C_1, C_2	0.0128	0.0141	$\mu F/km$
Zero	R_0	0.0293	0.1354	Ω/km
	L_0	2.6657	2.4874	mH/km
	C_0	0.0042	0.0046	$\mu F/km$

processing unit (NVIDIA GeForce GTX 2080 Ti), and it is carried out using a python-based version of TensorFlow (Google LLC) [25].

B. TRAINING HYPERPARAMETERS DETERMINATION

We varied the number of hidden layers when determining the optimal deep learning structure for saturation compensation. Five hidden layers were stacked during the pre-training described in section II. Fig. 6 shows that compensation was optimal with three or four hidden layers. However, due to computation burden and time efficiency, the suitable number of hidden layers is 3. Then, other optimal hyperparameters for the training framework afforded by BO are given in Table 3.

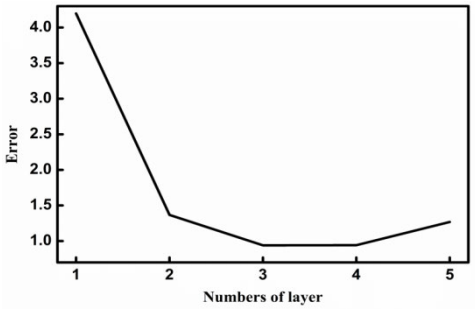

FIGURE 6. Error associated with different number of hidden layers.

TABLE 3. Optimal parameters for BDNN training.

Training parameters	Values		
	AE1	AE2	AE3
Hidden neuron	60	57	55
Learning rate	0.00078012	0.00035149	0.00028871
Batch size	192	256	192

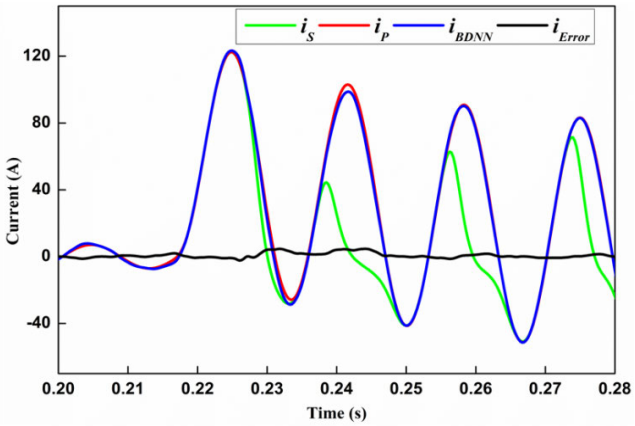

FIGURE 7. Compensation result on 345kV system with fault angle of 0° and remnant flux of 0%.

C. IMPACT OF BDNN FOR CT SATURATION

In this subsection, we investigate an impact analysis on 345kV overhead transmission system on the maximum DC offset occurs on the line. We consider scenarios with and without remnant flux. The result of CT saturation with and without remnant flux are illustrated in Fig. 7 and 8. Estimated outputs from the BDNN-, ANN-, and LES-based techniques are denoted as i_{BDNN}, i_{ANN}, and i_{LES}, respectively. We can observe in Fig. 7 that BDNN completely compensate for CT saturation throughout the whole cycle despite a slight and undershoot during the second fault cycle. Small oscillations are evident a fault, reflecting the use of moving-window algorithm and a sudden change in fault magnitude. However, this small error does not severely compromise the accuracy of the proposed BDNN. Another observation is that BDNN produces a slight undershoot in the second fault cycle after a fault occurs and it is acceptable in this study. By investigating

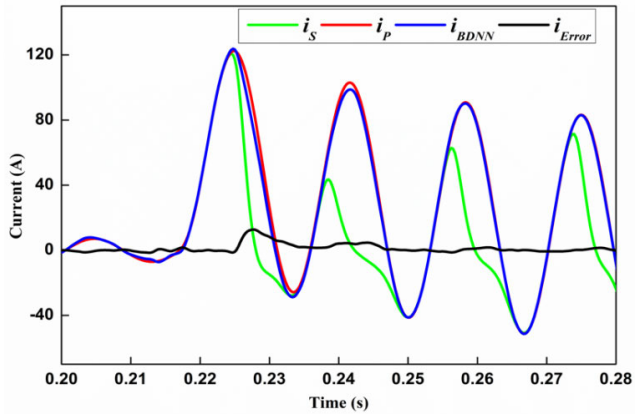

FIGURE 8. Compensation result on 345kV system with fault angle of 0° and remnant flux of 60%.

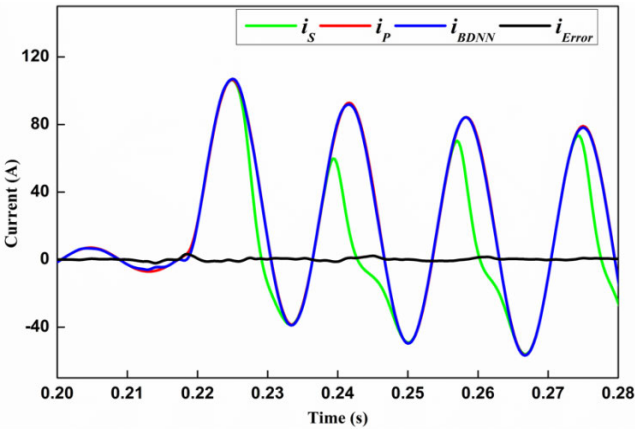

FIGURE 9. Compensation result on 345kV system with fault angle of 45° and remnant flux of 60%.

the error line (i_{Error}) of the proposed BDNN, the error line is shown where it yields the largest error between 0.2284s to 0.2472s. Furthermore, BDNN yields the error approximately 0.21 with remnant flux of 0%. Based on these observations, BDNN exhibits a great performance with less sensitivity to abovementioned issues and rapid convergence.

To further evaluate the efficiency during heavy saturation, the remnant flux is increased to 60% that produces extreme saturation on the fault angle of 0° as shown in Fig. 8. We observe in Fig. 8 that CT saturates in the first fault cycle that is approximately at 0.2219s. Similarly, BDNN estimates the correct magnitude of the unsaturated waveform. In addition, the use of moving-window technique on phasor estimation is insignificant as we discussed above. The error line generates a large swing when there is distortion in current from first half of fault cycle to the end of the second fault cycle and the error decay to nearly zero after the second fault cycle. Therefore, BDNN produces remarkable outputs even there is a heavy saturation occurred in the system.

Next, we investigate the performance of BDNN on saturation variation. It is vital for the proposed to estimate an accurate result at any given saturation variation occurred in the

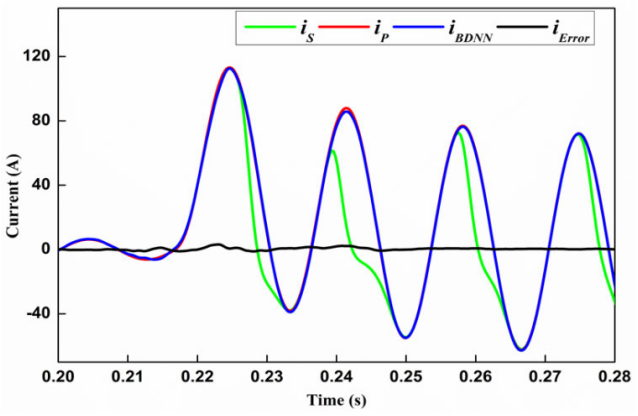

FIGURE 10. Compensation result on 154kV system with fault angle of 0° and remnant flux of 60%.

power systems. It is obvious that saturation severity decreases in proportional to the decrease of DC offset magnitude. Fig. 9 shows compensation results on fault angle of 45° with 60% remnant flux. In this case, the effect of moving-window algorithm yields more oscillation prior to fault. We can observe from these figures that BDNN also estimates magnitude with slight error after the fault. Likewise, the most noticeable error appears in the second fault cycle in which it produces an apparent undershoot on 345kV system.

It is crucial that the proposed BDNN can compensate for CT saturation in different voltage system. Fig. 10 presents the compensation result for another different voltage system with the severe saturation in the 154kV test system. As shown in Fig. 10, it is observed that BDNN can also estimate the correct magnitude in every cycle even it experiences the heavy saturation. Thus, we can conclude that the proposed BDNN performs well regardless voltage system because of input dataset normalization. Data normalization plays a significant role in our proposed BDNN algorithm as it reduces the influence of magnitude variation.

D. COMPARATIVE STUDY

To investigate the performance of ANN and LES filter, a comparative study is conducted. Fig. 11 shows the comparison results of ANN, LES, and BDNN on the most severe saturation having the fault angle of 0^0 and remnant flux of 60%.

As we observe in Fig. 11, the oscillation of ANN prior to fault is slightly apparent than BDNN and LES. As the fault occurs, ANN tends to produce less error than LES. As displayed in Fig. 10, LES apparently yields a few noticeable oscillations from the first to third fault cycle. It can be assumed that LES produces error when saturation is severe and the convergence of LES is achieved 3 cycles after a fault. Compared to LES, ANN requires 2 cycle to converge. Although ANN yields a similar output to the proposed BDNN, it produces a noticeable oscillation before a fault occurs and undershoot in the second fault cycle. Therefore, ANN and LES do not cope well with CT saturation compensation as they produce some ripples and overshoot in

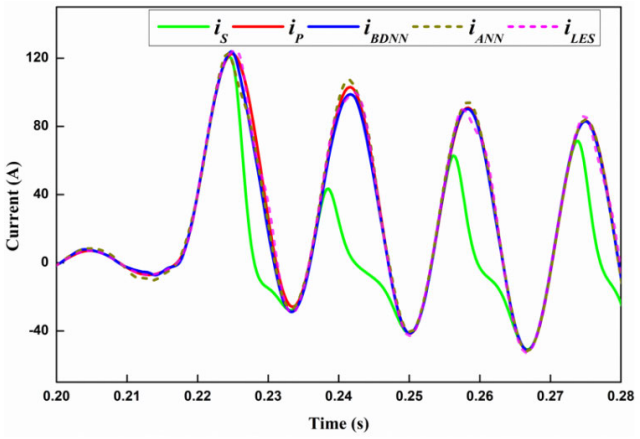

FIGURE 11. Compensation result on 345kV system with fault angle of 0° and remnant flux of 60% of ANN, and LES with BDNN.

TABLE 4. Mean error and standard deviation of BDNN, ANN, and LES.

Model	Fault angle (°)	Remnant flux (%)							
		0		20		40		60	
		μ	σ	μ	σ	μ	σ	μ	σ
154kV system									
LES [11]	0	0.50	0.56	0.53	0.56	0.59	0.67	0.63	0.76
	45	0.31	0.30	0.34	0.36	0.36	0.38	0.40	0.43
	90	0.11	0.25	0.15	0.33	0.19	0.38	0.20	0.31
ANN	0	0.41	0.54	0.45	0.69	0.45	0.69	0.47	0.72
	45	0.21	0.27	0.22	0.31	0.21	0.31	0.21	0.30
	90	0.15	0.19	0.20	0.18	0.20	0.29	0.21	0.27
BDNN	0	0.23	0.36	0.24	0.38	0.28	0.39	0.29	0.40
	45	0.16	0.22	0.17	0.23	0.18	0.24	0.17	0.27
	90	0.09	0.12	0.12	0.14	0.16	0.14	0.15	0.25
345kV system									
LES [11]	0	0.69	0.70	0.70	0.68	0.77	0.76	0.81	0.85
	45	0.66	0.69	0.66	0.67	0.71	0.70	0.76	0.80
	90	0.15	0.29	0.18	0.21	0.18	0.25	0.21	0.33
ANN	0	0.83	0.55	0.88	0.76	0.90	0.86	0.94	0.95
	45	0.32	0.53	0.31	0.50	0.31	0.48	0.34	0.57
	90	0.25	0.31	0.25	0.32	0.26	0.35	0.28	0.39
BDNN	0	0.38	0.59	0.38	0.41	0.43	0.56	0.39	0.41
	45	0.23	0.34	0.22	0.31	0.23	0.34	0.23	0.33
	90	0.08	0.18	0.11	0.21	0.15	0.25	0.17	0.26

some cases. Thus, the performance of BDNN gives better suppression on CT saturation when there is an appearance of CT saturation in the power systems. The comparison of other saturation cases is given in Table 4.

To evaluate the accuracy of CT saturation compensation for each algorithm, the estimation mean error (μ) and its standard deviation (σ) are computed and given in (13) and (14), respectively, where x_i is the reconstruction error between estimated and actual waveform. Table 4 summarizes estimation mean error and standard deviation of BDNN compared to ANN, and LES for several remnant flux and fault angle on a different voltage system. ANN generated mean error between 0.15 and 0.94 while LES returned error from 0.11 to 0.81. According to Table 4, BDNN reaches the largest error of 0.39 on 345kV system in the case of 0° fault angle with 60% remnant flux. ANN produced largest standard deviation of 0.95 on 345kV system in case of 0° fault angle with 60% remnant flux. BDNN achieved the smallest standard deviation between 0.12 and 0.59. Due to sudden change in magnitude after the fault occurrence and the use of moving-window technique, BDNN and ANN approximates the incorrect magnitude a cycle before a fault. However, BDNN shows less sensitive to the sudden increase after it experiences fault and provides quicker convergence than ANN as depicted in Figs. 7-11. ANN tends to produce a large swing one cycle before fault occurrence due to the moving-window effect. Thus, BDNN can compensate the effect of CT saturation quickly and give less ripple to the outputs under various fault magnitude. The effectiveness of BDNN and ANN can be implemented on other CTs that have similar characteristics. This fact was shown in some studies using intelligent methods for protective relays [3], [4].

$$\mu = \frac{1}{N}\sum_{i=1}^{N-1} x_i \tag{13}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N-1}(x_i - \mu)^2} \tag{14}$$

V. CONCLUSION

This paper presents a current transformer saturation compensation method comprising BO and DNN. DNN structure is established through stacking denoising autoencoders and backpropagation. By employing this method, the network appropriately obtains initial weights and bias for each hidden layer that reduces the computation complexity in the deeper structure and provide an easy method to build deep networks. Moreover, the utilization of SDAE is to make the model to suppress the noise effects in the real application without using the low-pass filter. BO perfectly optimizes the training hyperparameters which takes less time than other available optimizations for deep learning. The performance of BDNN is thoroughly evaluated on simulated data from PSCAD/EMTDC on the variation of saturation such as different fault angles, remnant flux, and power system level. The results show that BDNN can compensate CT saturation under various scenario regardless of fault types and fault current magnitude. Compared with ANN and LES, the reconstruction error of BDNN achieves the least error and its performance is relatively stable with different voltage system. Although the slight error of implementing moving-window algorithm is notably seen before a fault, this effect gives less influence on the phasor estimation. The limitation of the proposed BDNN is generalized to only some of the specific current transformer that share the similar characteristics to the CTs used during training phase. However, we intend to develop a compensation model using BDNN that can work for all kinds of current transformers. Our future work is to implement this proposed BDNN in the real time. The capability of CPU provided in AM572x, which is the hardware platform considered for our implementation in future, is 40GMAC per core (80G FLOP per core). The floating-point operation for the proposed BDNN is 1,201,104 that is sufficient for

CPU of real-time devices to calculate the neural network implementing on the real time.

REFERENCES

[1] A. Hargrave, M. J. Thompson, and B. Heilman, "Beyond the knee point: A practical guide to CT saturation," in *Proc. 71st Annu. Conf. Protective Relay Eng. (CPRE)*, College Station, TX, USA, Mar. 2018, pp. 1–23.

[2] D. C. Yu, J. C. Cummins, Z. Wang, H.-J. Yoon, and L. A. Kojovic, "Correction of current transformer distorted secondary currents due to saturation using artificial neural networks," *IEEE Trans. Power Del.*, vol. 16, no. 2, pp. 189–194, Apr. 2001.

[3] H. Khorashadi-Zadeh and M. Sanaye-Pasand, "Correction of saturated current transformers secondary current using ANNs," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 73–79, Jan. 2006.

[4] K. Erenturk, "ANFIS-based compensation algorithm for current-transformer saturation effects," *IEEE Trans. Power Del.*, vol. 24, no. 1, pp. 195–201, Jan. 2009.

[5] Y. C. Kang, U. J. Lim, S. H. Kang, and P. A. Crossley, "Compensation of the distortion in the secondary current caused by saturation and remanence in a CT," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1642–1649, Oct. 2004.

[6] S.-R. Nam, J.-Y. Park, S.-H. Kang, and M. Kezunovic, "Phasor estimation in the presence of DC offset and CT saturation," *IEEE Trans. Power Del.*, vol. 24, no. 4, pp. 1842–1849, Oct. 2009.

[7] E. Hajipour, M. Vakilian, and M. Sanaye-Pasand, "Current-transformer saturation compensation for transformer differential relays," *IEEE Trans. Power Del.*, vol. 30, no. 5, pp. 2293–2302, Oct. 2015.

[8] E. M. Esmail, N. I. Elkalashy, T. A. Kawady, A. M. I. Taalab, and M. Lehtonen, "Detection of partial saturation and waveform compensation of current transformers," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1620–1622, Jun. 2015.

[9] F. B. Ajaei, M. Sanaye-Pasand, M. Davarpanah, A. Rezaei-Zare, and R. Iravani, "Compensation of the current-transformer saturation effects for digital relays," *IEEE Trans. Power Del.*, vol. 26, no. 4, pp. 2531–2540, Oct. 2011.

[10] J. Pan, K. Vu, and Y. Hu, "An efficient compensation algorithm for current transformer saturation effects," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1623–1628, Oct. 2004.

[11] Y. Y. Hong and D. W. Wei, "Compensation of distorted secondary current caused by saturation and remanence in a current transformer," *IEEE Trans. Power Del.*, vol. 25, no. 1, pp. 47–54, Jan. 2010.

[12] F. Naseri, Z. Kazemi, E. Farjah, and T. Ghanbari, "Fast detection and compensation of current transformer saturation using extended Kalman filter," *IEEE Trans. Power Del.*, vol. 34, no. 3, pp. 1087–1097, Jun. 2019.

[13] A. Hooshyar and M. Sanaye-Pasand, "Accurate measurement of fault currents contaminated with decaying DC offset and CT saturation," *IEEE Trans. Power Del.*, vol. 27, no. 2, pp. 773–783, Apr. 2012.

[14] N. G. Chothani and B. R. Bhalja, "New algorithm for current transformer saturation detection and compensation based on derivatives of secondary currents and Newton's backward difference formulae," *IET Gener., Transmiss. Distrib.*, vol. 8, no. 5, pp. 841–850, May 2014.

[15] G. E. Hinton, S. Osindero, and Y. W. The, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1557, Jul. 2006.

[16] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2006, pp. 153–160.

[17] S. Key, V. Sok, S.-W. Lee, C.-S. Ko, S.-R. Nam, and N.-H. Lee, "Current transformer saturation compensation based on autoencoder and deep learning," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5.

[18] S.-B. Kim, V. Sok, S.-H. Kang, N.-H. Lee, and S.-R. Nam, "A study on deep neural network-based DC offset removal for phase estimation in power systems," *Energies*, vol. 12, no. 9, p. 1619, Apr. 2019.

[19] M. Ali, D.-H. Son, S.-H. Kang, and S.-R. Nam, "An accurate CT saturation classification using a deep learning approach based on unsupervised feature extraction and supervised fine-tuning strategy," *Energies*, vol. 10, no. 11, p. 1830, Nov. 2017.

[20] K. Chen, J. Hu, and J. He, "Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1748–1758, May 2018.

[21] S. Afrasiabi, M. Afrasiabi, B. Parang, and M. Mohammadi, "Integration of accelerated deep neural network into power transformer differential protection," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 865–876, Feb. 2020.

[22] S. Afrasiabi, M. Afrasiabi, B. Parang, M. Mohammadi, H. Samet, and T. Dragicevic, "Fast GRNN-based method for distinguishing inrush currents in power transformers," *IEEE Trans. Ind. Electron.*, early access, Sep. 8, 2021, doi: 10.1109/TIE.2021.3109535.

[23] S. Afrasiabi, M. Afrasiabi, B. Parang, and M. Mohammadi, "Designing a composite deep learning based differential protection scheme of power transformers," *Appl. Soft Comput.*, vol. 87, Feb. 2020, Art. no. 105975, doi: 10.1016/j.asoc.2019.105975.

[24] M. Afrasiabi, S. Afrasiabi, B. Parang, and M. Mohammadi, "Power transformers internal fault diagnosis based on deep convolutional neural networks," *J. Intell. Fuzzy Syst.*, vol. 37, no. 1, pp. 1165–1179, Jul. 2019, doi: 10.3233/JIFS-182615.

[25] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[26] IEEE Power System Relaying and Control Committee. (2003). *CT Saturation Theory and Calculator*. [Online]. Available: http//www.pes-psrc.org

[27] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 3371–3408, Dec. 2010.

[28] Tensorflow. (Sep. 2019). *Tensorflow API r1.15 Manual*. [Online]. Available: https://www.tensorflow.org/api_docs/python

[29] P. I. Frazier, "A tutorial on Bayesian optimization," 2018, *arXiv:1807.02811*.

[30] J. R. Lucas, P. G. McLaren, W. W. L. Keerthipala, and R. P. Jayasinghe, "Improved simulation models for current and voltage transformers in relay studies," *IEEE Trans. Power Del.*, vol. 7, no. 1, pp. 152–159, Jan. 1992.

[31] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, "Normalization techniques in training DNNs: Methodology, analysis, and application," in *Proc. Comput. Vis. Pattern Recognit.*, Sep. 2020, pp. 1–20. [Online]. Available: https://arxiv.org/pdf/2009.12836

SOPHEAP KEY received the bachelor's degree in electrical and electronic engineering from the Institute of Technology of Cambodia (ITC), Phnom Penh, Cambodia, in 2018. He is currently pursuing the M.S. and Ph.D. degrees in smart energy protection and automation with the Department of Electrical Engineering, Myongji University, South Korea. His current research interests include power quality, power system protection, and deep learning application on power systems.

SANG-HEE KANG (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Seoul National University, Seoul, South Korea, in 1985, 1987, and 1993, respectively.

He was a Visiting Fellow and a Visiting Scholar with the University of Bath, Bath, U.K., in 1991 and 1999, respectively. He has been with the Next-Generation Power Technology Center, South Korea, since 2001. He was an Honorary Academic Visitor with The University of Manchester, Manchester, U.K., in 2007. He is currently a Professor with Myongji University, Yongin, South Korea. His research interest includes to develop digital protection systems for power systems using digital-signal-processing techniques.

NAM-HO LEE (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Myongji University, South Korea, in 1998, 2001, and 2011, respectively. He is currently working as a Senior Researcher with the KEPCO Research Institute, Daejeon, South Korea. His research interests include IEC-61850-based substation automation and power restoration.

SOON-RYUL NAM (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Seoul National University, Seoul, South Korea, in 1996, 1998, and 2002, respectively.

He was with Hyosung Corporation, South Korea, from 2002 to 2005, and a Research Professor with Myongji University, Yongin, South Korea, from 2005 to 2007. He was a Postdoctoral Research Associate with Texas A&M University, College Station, in 2007, and an Assistant Professor with Chonnam National University, Gwangju, South Korea, from 2007 to 2009. He has been with Myongji University, since 2009. He is currently a Professor with Myongji University. His research interest includes the protection, control, and automation of power systems.