# Attention-GRU–Based Differential Protection for IBR-Interfaced Transformers

Ghansham Meghwar
Electrical and Computer Engineering
University of Illinois Chicago
Chicago, United States
gmeghw2@uic.edu

Lina He
Electrical and Computer Engineering
University of Illinois Chicago
Chicago, United States
lhe@uic.edu

***Abstract*—The increasing penetration of Inverter-Based Resources (IBRs) challenges conventional transformer differential protection schemes. The unique current and harmonic characteristics of IBRs can lead to differential relay misoperation during transient events such as magnetizing inrush. To address this, the paper proposes a novel differential protection scheme based on a Gated Recurrent Unit (GRU) network enhanced with a temporal attention mechanism. This architecture is designed to accurately classify transformer operating states by learning temporal dependencies from differential current waveforms and dynamically focusing on the most discriminative time steps to distinguish between internal faults, external faults, and inrush currents. The proposed model is evaluated on a comprehensive dataset generated from PSCAD simulations of a modified IEEE 39-bus system integrated with a photovoltaic (PV) plant.**

***Index Terms*--Differential protection, fault classification, Gated Recurrent Unit (GRU), Inverter-Based Resources (IBR), temporal attention mechanism.**

## I. Introduction

The accelerating integration of Inverter-Based Resources (IBRs), such as large-scale photovoltaic (PV) plants, into modern power grids fundamentally changes system dynamics and poses significant challenges to conventional protection philosophies [1],[2],[3]. Differential protection, a critical scheme for power transformer protection, is particularly affected by this transition [4]. Traditional differential relays are designed based on the predictable fault current characteristics of synchronous generators, which produce high magnitude, sustained fault currents. In contrast, IBRs exhibit current-limited behavior and produce distinct harmonic and sequence component profiles during transient events [5]. This disparity can lead to the misoperation of protection relays, most notably false tripping during non-fault conditions, such as magnetic inrush, a phenomenon identified as a key challenge in recent industry analyses of IBR-dominated systems [6].

This paper is based upon work supported by the U.S. Department of Energy (DOE) Solar Energy Technologies Office (SETO) and Wind Energy Technologies Office (WETO) under the Solar and Wind Grid Services and Reliability Demonstration Funding Program, Award Number: DE-EE0010656. This research was conducted at the University of Illinois Chicago.

The primary mechanism behind this challenge lies in the harmonic content of the differential currents. Conventional relays rely on second and fifth harmonic blocking to distinguish internal faults from magnetizing inrush [7],[8]. However, the harmonic signature of inrush currents in transformers fed by grid-following (GFL) inverters differs from that of synchronous sources, which may cause the blocking logic to fail. While operational strategies, such as adjusting harmonic blocking thresholds, have been explored to mitigate false trips, these modifications can create a trade-off, potentially reducing the relay's sensitivity to legitimate internal faults. This highlights a critical need for more sophisticated and adaptive protection schemes that can accurately distinguish between fault and non-fault transients in the complex environment of an IBR-integrated grid [9],[10].

In response to the need for faster and more reliable differential protection, researchers have increasingly turned to data-driven and deep learning techniques. A significant portion of this research has focused on the application of Convolutional Neural Networks (CNNs) due to their proficiency in extracting salient features from raw current signals. These approaches include using an Attention-module embedded Fully Convolutional Network (A-FCN) to handle variable-length signals and focus on non-inrush current waveforms [11], as well as developing accelerated CNNs designed to minimize computational overhead for real-time fault detection [12]. Other methods integrate signal processing with neural networks, such as the Signal Localized CNN (SLCNN), which performs a wavelet transform to capture non-stationary features [13]. More advanced architectures have utilized adversarial training to focus the network on discriminative, unsaturated features [14]. While these CNN-based methods are powerful, their primary strength in extracting localized features may not fully capture the sequential, time-dependent nature of transient phenomena in transformers.

979-8-3315-8138-1

To better address the temporal dynamics of differential currents, researchers have investigated recurrent and hybrid architectures. A notable example is a composite deep learning model that integrates a CNN with a Light-Gated Recurrent Unit (LGRU), designed to simultaneously learn spatial and temporal dependencies from the data. This hybrid approach demonstrates the recognized value of incorporating sequential modeling for improving classification accuracy [15]. Beyond deep learning, other machine learning paradigms have been applied, such as a scheme combining Intrinsic Time-Scale Decomposition (ITD) with an Ensemble of Bagged Trees (EBT) classifier to perform feature extraction in the time domain [16]. Furthermore, deep neural networks (DNNs) built upon stacked autoencoders have been utilized for automatic feature learning to create a threshold-less protection scheme [17]. Another distinct application involves a Bayesian Deep Neural Network (BDNN) that focuses specifically on compensating for current transformer (CT) saturation, rather than fault classification, highlighting the versatility of neural networks in the protection domain [18].

While the existing literature presents many advanced solutions, they often rely on complex signal pre-processing or computationally intensive hybrid structures. In contrast, this paper proposes a more concise and efficient end-to-end solution: a Gated Recurrent Unit (GRU) network enhanced with a temporal attention mechanism. The GRU architecture is selected for its effectiveness in modeling sequential data, making it inherently suitable for time-series current waveforms. The key contribution is the integration of the attention mechanism, which enables the model to dynamically focus on the most discriminative steps to distinguish between challenging fault and inrush conditions. To ensure modern relevance, the proposed method is validated on a power transformer connecting a PV plant to the IEEE 39-bus grid, directly addressing the emerging protection challenges in systems with high IBR penetration.

The remainder of this paper details the proposed temporal attention-based GRU architecture in Section II. Section III describes the modified IEEE 39-bus simulation environment and the data generation methodology. Section IV presents the model's performance evaluation and a discussion of the results, followed by Section V, which concludes the paper with a summary of findings and suggestions for future research.

## II. Proposed Temporal Attention-Based GRU Architecture

### A. Input Signal Processing and Model Workflow

The proposed protection scheme is designed to learn from the dynamic behavior of the transformer's three-phase differential currents. The workflow begins with the primary $I_P$ and secondary $I_S$ currents measured at the transformer terminals for each phase. From these, the three-phase differential current vector, $[I_{dA}, I_{dB}, I_{dC}]$, is calculated. To prepare this signal for the deep learning model, the continuous current waveforms are first digitized at a sampling rate of 6.72 kHz, corresponding to 112 samples per 60 Hz cycle. For each simulated event, a data window of approximately half a cycle is extracted. This process yields a fixed-length sequence of 57 time-step samples for each event. Before this sequence is fed into the network, it undergoes a crucial preprocessing step where it is normalized using the Standard Scaler transformation, which is fit only on the training data to prevent data leakage. The final processed input, a tensor of shape 57×3, serves as the input vector $X_T$ for the proposed GRU model. The specific architecture designed to perform this classification is detailed in the following section.

### B. Model Architecture

The proposed model for transformer differential protection leverages a GRU network enhanced with a temporal attention mechanism to accurately detect and classify fault conditions [19],[20]. The overall architecture, depicted in Fig. 1, consists of three main stages: a GRU layer for feature extraction, a temporal attention layer, and an output layer for final classification.

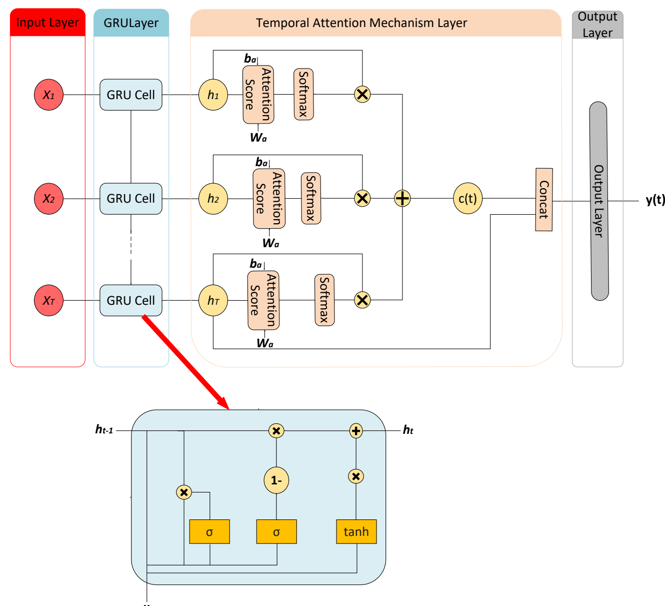


Fig. 1. Architecture of the proposed temporal attention-based GRU

The initial stage of the model consists of a GRU layer that processes the input sequence of normalized differential-current measurements, $\mathbf{X}_T = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T)$, where $T$ is the total number of time steps. For each time step $t$, the GRU cell updates its hidden state $\mathbf{h}_t$ by employing a gating mechanism. This mechanism utilizes an update gate, $\mathbf{z}_t$, and a reset gate, $\mathbf{r}_t$, to manage the flow of information across time steps. The operations are defined by the following equations [19],[20]:

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (1)$$

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (2)$$

Here, $x_t$ is the input at the time step $t$, $h_{t-1}$ is the hidden state from the previous time step, and $\sigma$ is the sigmoid activation function. The terms $W_{xz}$, $W_{hz}$, $W_{xr}$, $W_{hr}$ are weight matrices, while $b_z$ and $b_r$ are bias vectors. A candidate hidden state, $\tilde{h}_t$, is then computed based on the current input and the reset portion of the previous hidden state:

$$\tilde{h}_t = \phi(W_{nx}x_t + W_{nh}(r_t \odot h_{t-1}) + b_h) \quad (3)$$

where $\odot$ denotes the elementwise Hadamard product and $\phi$ is the hyperbolic tangent activation function. Finally, the update gate $z_t$ linearly interpolates between the previous hidden state $h_{t-1}$ and the candidate's hidden state $\tilde{h}_t$ to produce the final hidden state $h_t$ for the current time step:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

This process is repeated across all time steps, resulting in a sequence of hidden states $H = (h_1, h_2, \dots, h_T)$ that encode the temporal features of the input-current sequence.

Following the GRU layer, a temporal attention mechanism is applied to the sequence of hidden states. This layer enables the model to dynamically assign varying levels of importance to the features extracted at different time steps. By focusing on the most discriminative moments in the sequence, the model can enhance its sensitivity to fault-induced transients. First, an attention score, $e_t$, is calculated for each hidden state $h_t$ to quantify its relevance.

$$e^t = v_a^{\top} \phi(W_a h_t + b_a) \quad (5)$$

where $W_a$ and $b_a$ are learnable weight and bias parameters of the attention network, respectively, and $v_a^{\top}$ is a learnable projection vector. These scores are then normalized into attention weights, $\alpha_t$, using the softmax function.

$$\alpha_t = \frac{\exp(e^t)}{\sum_{k=1}^{T} \exp(e_k)} \quad (6)$$

The resulting attention weights are used to compute a context vector, $c(t)$, which is the weighted sum of all hidden states. This vector serves as a comprehensive and focused summary of the entire input sequence:

$$c(t) = \sum_{t=1}^{T} \alpha_t \, h_t \quad (7)$$

For the final classification stage, the information-rich context vector $c(t)$ is concatenated with the final hidden state of the GRU sequence, $h_T$. This combined representation, containing both a summary of the most important sequence features and the most recent state information, is then passed to a fully connected output layer. The final layer produces the prediction, $y(t)$, which indicates the operational state of the transformer.

## *C. Importance of Temporal Attention Mechanism*

A primary challenge in transformer differential protection is the high degree of similarity that can exist between magnetizing inrush and certain internal faults, an issue illustrated by the waveforms in Fig. 2. Conventional numerical relays typically struggle with such cases because their logic relies heavily on the magnitude of the differential current and its second-harmonic content [21], often analyzed on a per-phase basis. This approach is vulnerable because it may fail to capture the inherently asymmetrical nature of inrush currents across the three phases, a key distinguishing characteristic compared to many symmetrical internal faults.

While a standard GRU network improves upon this by processing the complete three-phase current sequences, providing it with a richer temporal context that can capture these asymmetries, it is not without its own limitations. The network processes the entire sequence, and its final hidden state can be overly influenced by dominant but less discriminative features, such as the large, similar peak magnitude of approximately −85A shown in Fig. 2, potentially overlooking the subtle but decisive cues present elsewhere in the waveform.

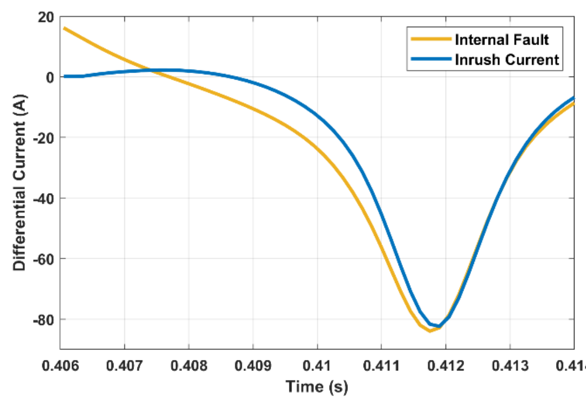


Fig. 2. Differential current waveforms for an internal fault and an inrush event

The proposed attention layer addresses these limitations by dynamically weighting the differential current sequences. By learning to focus its attention on discriminative features, such as the slightly steeper initial descent and the sharper curvature at the peak of the internal fault waveform, the model can accurately distinguish the fault from the inrush event. This selective focus prevents critical, low-magnitude information from being overshadowed by high-magnitude similarities, demonstrating a level of detailed analysis that is essential for robust protection in modern power systems.

# III. Simulated Power System Network

The power system model utilized for this study is a modified version of the IEEE 39-bus New England test system, a standard benchmark for power system stability and protection studies. The single-line diagram of the modified system is depicted in Fig. 3. A utility-scale IBR is interconnected at Bus 23 through a dedicated step-up transformer. The IBR is modeled as a PV-fed, GFL voltage-source converter (VSC) with a phase-locked, inner dq current control with feed-forward decoupling, and outer active/reactive power regulators. The model includes current-limiting and fault ride-through (FRT) constraint logic; the converter is interfaced to the low voltage (LV) winding of the transformer through an LC filter at the point of common coupling (PCC), and its DC link is supplied by a PV array with Maximum Power Point Tracking (MPPT).

979-8-3315-8138-1

The step-up transformer connecting the IBR's PCC to Bus 23 is the primary focus of this study. This critical component is a 300MVA, 345/34.5 kV, 60 Hz transformer with a Yg/Yg connection, modeled with copper losses of 0.005 p.u., iron losses of 0.0001 p.u., and a leakage reactance of 0.1 p.u. The data for training and evaluating the proposed deep learning-based differential protection relay are sourced from the primary and secondary currents of this transformer. This configuration provides a realistic testbed for analyzing the challenges and dynamic behaviors introduced by IBRs in a conventional protection scheme.

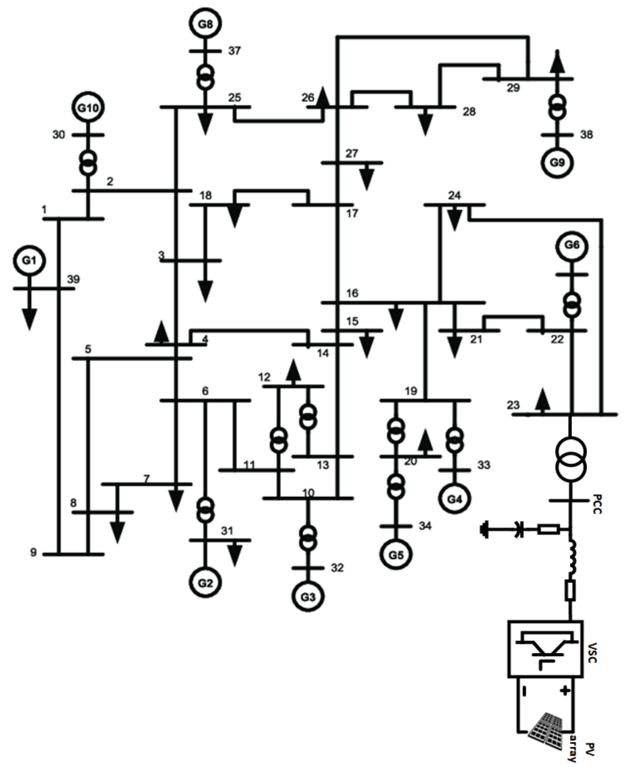


Fig. 3. Single-line diagram of the modified IEEE 39-bus test system with the interconnection of the PV-fed IBR and the step-up transformer at Bus 23.

## A. Data Generation

As detailed in Table I, a comprehensive dataset was generated by simulating a wide range of operating conditions to ensure the algorithm's robustness and accuracy. The simulations cover three main categories: magnetizing inrush, internal faults, and external faults. For these events, a diverse set of parameters, including fault type (line to ground (LG), line to line (LL), line to line to ground (LLG), triple line (LLL) and triple line to ground (LLLG), fault class (turn to ground (TG), turn to turn (TT)) resistance, inception angle, CT burden, and remanent flux, was varied to create an exhaustive dataset. This provides a rigorous basis for validating the performance of the protection scheme under challenging system conditions.

Following the generation of the complete dataset from all simulation cases, the samples were partitioned to facilitate robust model training and unbiased evaluation. The data were randomly split into a training set (70% of the total samples), a validation set (15%), and a test set (15%). The training set was used exclusively for learning the model parameters. The validation set was used during the training process to monitor for overfitting and to tune hyperparameters. Finally, the test set, which the model never sees during training, was reserved for the final performance assessment reported in Section IV, ensuring an objective evaluation of the model's generalization capability.

TABLE I. SIMULATION CASES

| Case | Factor | Levels (values) | Total samples |
|---|---|---|---|
| Inrush | CT burden | 2, 5, 10, 20 Ω | 4×4×8×7 = 896 |
| | CT remanent flux density | 0, 0.5, 1, 1.5 T | |
| | Transformer remanent flux | 0.1 p.u. to 0.8 p.u. (step of 0.1) | |
| | Transformer energization angle | 0° to 270° (step 45°) | |
| Internal fault | Fault class | TG, TT | 2×4×2×3×7 = 336 |
| | Faulted winding % | 20% to 80% (Step 20%) | |
| | Winding side | HV, LV | |
| | Phase | A, B, C | |
| | Inception angle | 0° to 270° (step 45°) | |
| Internal fault | Fault type | LG, LL, LLL, LLLG | 4×5×2×7 = 280 |
| | Fault resistance | 0.01, 0.1, 0.5, 1, 10 Ω | |
| | Winding side | HV, LV | |
| | Fault inception angle | 0° to 270° (step 45°) | |
| External fault | Fault type | LG, LL, LLL, LLLG | 4×5×7 = 140 |
| | Fault resistance | 0.01, 0.1, 0.5, 1, 10 Ω | |
| | Fault inception angle | 0° to 270° (step 45°) | |

# IV. PERFORMANCE EVALUATION AND DISCUSSION

## A. Experimental Setup

To ensure a rigorous and fair comparison, the proposed model was evaluated against three established baseline architectures: a standard GRU, a Long Short-Term Memory (LSTM) network, and a CNN. All models were implemented in PyTorch and trained under a standardized protocol to isolate performance differences attributable to their architectural merits. This protocol consisted of training for 256 epochs with a batch size of 32 and using the Adam optimizer with a Binary Cross-Entropy (BCE) loss function.

The proposed architecture is specifically configured as a two-layer stacked GRU with a hidden size of 128 units, followed by a multi-head attention mechanism. To mitigate overfitting, a dropout rate of 0.4 was applied to both the GRU and the subsequent fully connected layers. The Adam optimizer for this model was further specified with a learning rate of 0.0006 and a weight decay of 0.0005.

## B. Comparative Results

The comparative results, summarized in Table II, demonstrate the superior performance of the proposed temporal-attention-based model. It achieves the highest overall accuracy of 99.18% and an F1-score of 98.89%. Notably, all tested methods achieved a precision of 100.00%, indicating that no false trips (false positives) occurred in the test set. Given the perfect precision, the key performance differentiators are recall, which reflects the model's ability to correctly identify true fault events, and the F1-score. The proposed method delivers a recall of 97.80%, outperforming the CNN (97.12%), GRU (95.88%), and LSTM (95.68%). This improvement in recall directly translates to the leading F1-score.

TABLE II TEST SET PERFORMANCE COMPARISON

| Metric | Proposed algorithm | Plain GRU | CNN | LSTM |
|---|---|---|---|---|
| Accuracy | 99.18% | 98.36% | 98.77% | 98.15% |
| Precision | 100.00% | 100.00% | 100.00% | 100.00% |
| Recall | 97.80% | 95.88% | 97.12% | 95.68% |
| F1 score | 98.89% | 97.89% | 98.54% | 97.79% |

### C. Discussion

The significant performance gain is directly attributable to the temporal attention module. This result provides strong evidence for the hypothesis presented in Section II-C, which argues that the ability to focus on subtle, discriminative features is critical for distinguishing between challenging inrush and internal fault waveforms. In contrast to conventional recurrent models, such as GRU and LSTM, which compress the entire sequence into a single terminal hidden state, or CNNs that rely on a fixed pooling operation, the attention mechanism dynamically assigns higher weights to the most diagnostic instances within the time-series window. This capability enables the model to focus on critical data points, resulting in fewer missed internal-fault detections. Ultimately, the proposed architecture enhances dependability (higher recall) while maintaining perfect security (no false positives). This is a highly desirable property for critical applications such as transformer differential protection, especially within inverter-rich networks where complex fault signatures are prevalent.

## V. CONCLUSION

This paper addresses the challenge of transformer differential protection in IBR-penetrated grids by proposing a novel scheme that utilizes a GRU with a temporal attention mechanism. Validated on a modified IEEE 39-bus system with an integrated PV plant, the model demonstrated superior performance over traditional deep learning architectures, including standard GRU, LSTM, and CNN models, achieving an overall accuracy of 99.18% and an F1-score of 98.89%. The key finding is the model’s ability to achieve a perfect balance between protection, security, and dependability.

This success is attributed to the attention layer, which dynamically focuses on the most diagnostically significant moments in the current signal to distinguish internal faults from events like magnetic inrush. The results confirm that the GRU-attention architecture provides a robust, accurate, and reliable solution for modern power grids. Future work could explore the real-time implementation of this approach on hardware-in-the-loop platforms and its performance across a broader range of IBR models.

## REFERENCES


[1] M.-A. Nasr and A. Hooshyar, "Power swing in systems with inverter-based resources—part II: impact on protection systems," *IEEE Trans. Power Delivery*, vol. 39, pp. 1903-1917, Jun. 2024.

[2] G. Meghwar, S. Yu, and L. He, "An enhanced fault detection strategy based on the transient voltage in multi-terminal HVDC grid," in *Proc. 2024 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1-5.

[3] A. A. Memon, A. Kumar, G. Sham, and U. Yasir, "Automated protection of power transformers at distribution grid connected to green energy sources," in *Proc. 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, pp. 1-4.

[4] S. Afrasiabi, B. Behdani, M. Afrasiabi, M. Mohammadi, Y. Liu, and M. Gheisari, "A comparative analysis of artificial intelligence for power transformer differential protection," in Proc. *2021 IEEE International Conf. on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, pp. 1-6.

[5] B. Ahmadzadeh-Shooshtari and A. Rezaei-Zare, "Advanced transformer differential protection under GIC conditions," *IEEE Trans. Power Delivery*, vol. 37, pp. 1433-1444, Jun. 2022.

[6] S. Grogan and B. Badrzadeh, "Topic 5 stage 4 restoration and black start," CSIRO, Australia, Tech. Rep., 2025.

[7] E. Hajipour, M. Vakilian, and M. Sanaye-Pasand, "Current-transformer saturation compensation for transformer differential relays," *IEEE Trans. Power Delivery*, vol. 30, pp. 2293-2302, Oct. 2015.

[8] G. Meghwar, D. Weicheng, M. Bawazeer, and J. Xu, "Reciprocating connection of dual active bridge DC/DC converter with PV integration," in *Proc. 2020 5th International Conf. on Power and Renewable Energy (ICPRE),* pp. 566-571.

[9] X. Chen, A. Haddadi, E. Farantatos, I. Kocar, and S. Bu, "Power swing in systems with varying penetration of grid-forming IBRs: protection and dynamics," *Prot. Control Mod. Power Syst.*, vol. 10, pp. 116-129, Jul. 2025.

[10] M. Bawazeer, D. Weicheng, G. Meghwar, and J. Xu, "Reciprocating bus in FREEDM system with solar farm integration," in *Proc. 2020 5th International Conf. on Power and Renewable Energy (ICPRE)*, pp. 27-32.

[11] Z. Lin, S. Yang, Y. Zhang, Z. Hao, and B. Zhang, "A data-aided power transformer differential protection without inrush blocking module," *IEEE Trans. Power Delivery*, vol. 38, pp. 2000-2010, Jun. 2023.

[12] 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, pp. 865-876, Feb. 2020.

[13] S. K. Murugan, S. P. Simon, and R. R. Eapen, "A novel signal localized convolution neural network for power transformer differential protection," *IEEE Trans. Power Delivery*, vol. 37, pp. 1242-1251, Apr. 2022.

[14] A. He, Z. Jiao, Z. Li, and Y. Liang, "Discrimination between internal faults and inrush currents in power transformers based on the discriminative-feature-focused CNN," *Electr. Power Syst. Res.*, vol. 223, p. 109701, Oct. 2023.

[15] 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, p. 105975, Feb. 2020.

[16] Z. Moravej, A. Ebrahimi, M. Pazoki, and M. Barati, "Time domain differential protection scheme applied to power transformers," *Int. J. Electr. Power Energy Syst.*, vol. 154, p. 109465, Dec. 2023.

[17] S. Key, G.-W. Son, and S.-R. Nam, "Deep learning-based algorithm for internal fault detection of power transformers during inrush current at distribution substations," *Energies*, vol. 17, p. 963, Feb. 2024.

[18] S. Key, S. -H. Kang, N. -H. Lee and S. -R. Nam, "Bayesian deep neural network to compensate for current transformer saturation," in *IEEE Access*, vol. 9, pp. 154731-154739, 2021, doi: 10.1109/ACCESS.2021.3127542.

[19] K. Cho, et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," unpublished, Jun. 2014. [Online]. Available: https://arxiv.org/abs/1406.1078

[20] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," unpublished, Sep. 2014. [Online]. Available: https://arxiv.org/abs/1409.0473

[21] S. H. Horowitz, A. G. Phadke, and C. F. Henville, *Power System Relaying*. New York: Wiley, 2022.