ISSN NO: 9726-001X
Volume 13 Issue 02 2025

Adversarial Attacks on Deep Learning-Based Supervisory Protection Systems for Wide AreaMonitoring in Smart Grids: A PMU Data Perspective

B.Hamsaskanditha[1], CH.Hareesh [2], K.Srinija [3], Mr.G.Poshamallu [4]
[1,2,3] UG Scholar, Department of ECE, St. Martin's Engineering College, Secunderabad, Telangana,India-500100
[4]Assistant Professor, Department of ECE, St. Martins Engineering College, Secunderabad, Telangana,India-500100
skanditha17@gmail.com

Abstract:

The modern smart grid has undergone a significant transformation by incorporating advanced monitoring systems that leverage Phasor Measurement Unit (PMU) data for real time state estimation and fault detection. Traditional supervisory protection systems rely on rule-based algorithms and statistical methods to process PMU data; however, these approaches often struggle with the increasing complexity and dynamic nature of modern power networks. Their limitations are further com- pounded by the emergence of adversarial attacks, which can subtly manipulate sensor data and lead to misinterpretations by deep learning models. This presents a critical challenge to grid stability and security. In traditional systems, PMU devices are connected to a centralized Data Acquisition System (DAS) that feeds into SCADA and state estimation modules. Although effective under normal conditions, these systems are limited by static thresholding techniques, delayed response times, and an inability to adapt to non-linear grid behaviours. As adversarial techniques evolve, these legacy systems lack the robustness needed to detect and mitigate sophisticated cyber-physical threats, resulting in increased risks of grid failures and economic losses. The need for a more resilient and adaptive system is clear. In response, the proposed methodology integrates a hybrid model that combines a Convolutional Neural Network (CNN) for robust feature extrac- tion with a Random Forest Classifier (RFC) for reliable decision making. This approach leverages the hierarchical feature learning capabilities of deep learning to process complex PMU data, while the ensemble strength of RFC enhances predictive accuracy and robustness against adversarial perturbations. The significance of this proposed system lies in its potential to revolutionize grid monitoring by offering improved accuracy, faster response times, and enhanced security. By addressing both the vulnerabilities of traditional systems and the emerging threats of adversarial attacks, the hybrid model ensures a more secure and stable operation of modern smart grids, paving the way for future advancements in critical infrastructure protection.

***Index Terms*—Adversarial attacks, deep learning, smart grid, CNN, PMU data, security.**

I. INTRODUCTION

The evolution of smart grids has brought unprecedented efficiency and resilience to modern power systems, enabling real-time monitoring, rapid fault detection, and seamless integration of renewable energy sources. At the heart of this transformation lies Phasor Measurement Units (PMUs), which provide highly accurate, time-synchronized data essential for grid stability. However, as power systems become increasingly dependent on deep learning-based supervisory protection mechanisms, they also become prime targets for adversarial attacks. Cybercriminals can subtly manipulate PMU data, deceiving intelligent models into misclassifying critical grid conditions, potentially leading to catastrophic blackouts, severe economic losses, and national security threats. Existing rule-based protection frameworks fail to counter these evolving threats, necessitating the development of a more robust and adaptive security solution. This paper proposes a groundbreaking hybrid architecture that synergizes the feature-extraction capabilities of Convolutional Neural Networks (CNNs) with the decision-making prowess of a Random Forest Classifier (RFC). By leveraging deep learning's ability to analyze complex PMU data and integrating ensemble-based classifiers for enhanced robustness, this research pioneers a next-generation defense mechanism for modern smart grids. The proposed model not only enhances the accuracy and reliability of grid monitoring but also fortifies critical infrastructure against the ever-growing landscape of cyber-physical threats.

II. EXISTING SYSTEM

Conventional wide-area monitoring and protection systems rely on Phasor Measurement Units (PMUs) to collect highspeed, time-synchronized data, which is then processed using traditional state estimation algorithms and rule-based protection mechanisms. These systems leverage Supervisory Control and Data Acquisition (SCADA) frameworks to detect anomalies and ensure grid stability. However, their reliance on predefined thresholds and deterministic models makes them increasingly ineffective in handling the complexities of modern power networks. As grids become more dynamic with the integration of renewable energy sources, conventional methods struggle to adapt to fluctuating conditions, non-linear behaviors, and emerging cyber threats. Furthermore, these legacy systems are not designed to process the massive influx of real-time data generated by an expanding network of PMUs, leading to computational bottlenecks and delayed decision-making. Most critically, traditional protection mechanisms lack robust cybersecurity measures, making them highly vulnerable to adversarial attacks that manipulate sensor data to deceive fault detection systems. With the growing sophistication of cyber-physical threats, the limitations of conventional grid monitoring necessitate a shift towards intelligent, learning-based solutions capable of real-time anomaly detection, adaptive decision-making, and enhanced resilience against cyber intrusions. Fig. 1 illustrates the traditional architecture.

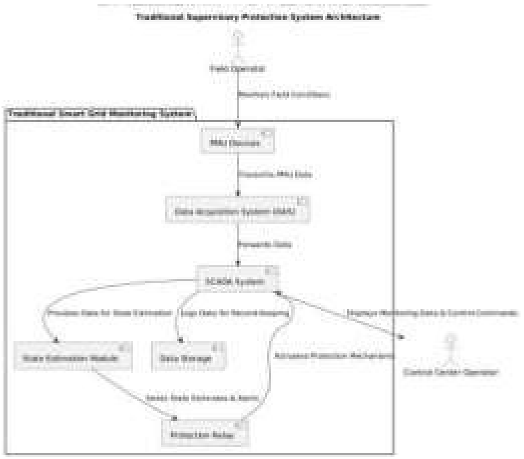

Traditional Supervisory Protection System Architecture
Field Operator
Monitors Field Conditions
Traditional Smart Grid Monitoring System
PMU Devices
Transmits PMU Data
Data Acquisition System (DAS)
Forwards Data
SCADA System
Provides Data for State Estimation
Logs Data for Record Keeping
Displays Monitoring Data & Control Commands
State Estimation Module
Data Storage
Activates Protection Mechanisms
Control Center Operator
Sends State Estimates & Alerts
Protection Relay

Fig. 1. Traditional smart grid monitoring architecture.

III. Proposed System

The proposed system introduces a robust and intelligent framework for monitoring and protecting smart grids against adversarial attacks by integrating deep learning and ensemble-based classification techniques. Unlike traditional rule-based supervisory protection mechanisms, this approach leverages the feature extraction capabilities of Convolutional Neural Networks (CNNs) and the decision-making efficiency of a Random Forest Classifier (RFC). The system is designed as an interactive desktop application with a user-friendly graph- ical interface built using Python's Tkinter library, allowing users to seamlessly upload datasets, preprocess data, train models, visualize results, and make real-time predictions.

The methodology follows a structured workflow, beginning with data preprocessing, class balancing using SMOTE, and image resizing to ensure consistency in model input. The hybrid approach enhances both the accuracy and robustness of anomaly detection by utilizing CNNs for deep feature extraction and RFC for final classification. This model significantly outperforms traditional classifiers like Gaussian Na¨ıve Bayes (GNB) and standalone RFC, demonstrating superior resilience against adversarial perturbations while maintaining real-time processing efficiency. By integrating real-time visualization of classification results, confusion matrices, and performance metrics, the proposed system ensures a more secure, adaptive, and intelligent supervisory protection system for modern smart grids.

IV. Model Training

A. Gaussian Na¨ıve Bayes Classifier (GNB)

Gaussian Na¨ıve Bayes is a probabilistic classifier that operates based on Bayes' theorem, assuming that all features are conditionally independent. It is chosen as a baseline model due to its simplicity, computational efficiency, and ability to provide quick results. However, its major drawback lies in its assumption of feature independence, making it less effective in handling complex, high-dimensional PMU data.

B. Random Forest Classifier (RFC)

The Random Forest Classifier is an ensemble learning algorithm that constructs multiple decision trees and merges their outputs to enhance predictive accuracy and reduce overfitting. It is particularly useful in handling high-dimensional datasets and provides better generalization capabilities than single decision trees. By training on PMU data, RFC improves classification robustness and ensures reliable anomaly detection.

C. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is employed to extract hierarchical feature representations from PMU image data. The model consists of multiple convolutional layers for detecting spatial dependencies, pooling layers for dimensionality reduction, and fully connected layers for classification. CNNs excel in identifying subtle patterns in PMU data, making them highly effective in distinguishing between normal operations and adversarial attacks.

D. Hybrid CNN-RFC Model

The final proposed model integrates CNN as a feature extractor with RFC as the classifier, combining the strengths of deep learning and ensemble methods. After training the CNN, the final classification layer is removed, and the extracted feature maps are fed into the RFC for decision-making. This hybrid approach enhances the system's ability to de- tect adversarial manipulations while leveraging RFC's robust classification capabilities, significantly improving accuracy, adaptability, and resilience against cyber-physical threats.

By combining these models, the proposed system ensures enhanced security, real-time anomaly detection, and improved grid stability, making it a cutting-edge solution for modern smart grid monitoring. Fig. 2 depicts the proposed system.

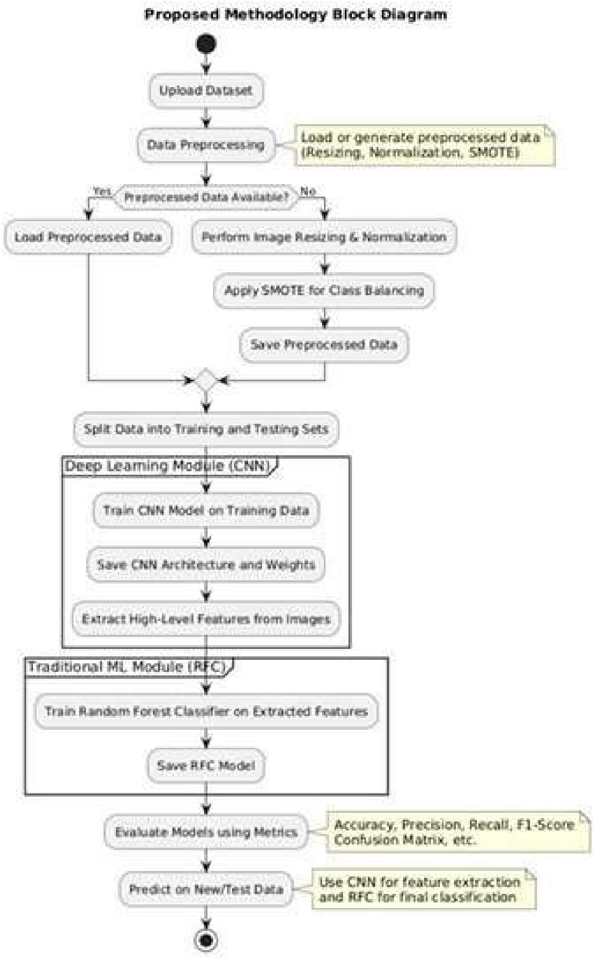

Proposed Methodology Block Diagram
Upload Dataset
Data Preprocessing
Load or generate preprocessed data (Resizing, Normalization, SMOTE)
Yes
Preprocessed Data Available?
No
Load Preprocessed Data
Perform Image Resizing & Normalization
Apply SMOTE for Class Balancing
Save Preprocessed Data
Split Data into Training and Testing Sets
Deep Learning Module (CNN)
Train CNN Model on Training Data
Save CNN Architecture and Weights
Extract High-Level Features from Images
Traditional ML Module (RFC)
Train Random Forest Classifier on Extracted Features
Save RFC Model
Evaluate Models using Metrics
Accuracy, Precision, Recall, F1-Score Confusion Matrix, etc.
Predict on New/Test Data
Use CNN for feature extraction and RFC for final classification

Fig. 2. Proposed hybrid CNN-RFC model.

V. Results and Discussion

A. Implementation Description

The proposed project is a comprehensive application that integrates data preprocessing, model training, performance evaluation, and real-time prediction functionalities through an intuitive graphical user interface (GUI). Built using Python's Tkinter library, the GUI allows users to interact seamlessly with the system while leveraging machine learning and deep learning libraries such as scikit-learn, imbalanced-learn, TensorFlow/Keras, and OpenCV for data processing and classification. The system comprises multiple modules, including data handling, model training, evaluation, and visualization, ensuring a streamlined workflow from dataset selection to final classification.

B. Module Imports and Global Variables

The application begins by importing essential libraries that handle GUI interactions, numerical computations, and machine learning functionalities. Tkinter and filedialog modules facilitate dataset selection, while NumPy, Pandas, Seaborn, and Matplotlib support data manipulation and visualization. The machine learning module integrates traditional classifiers like Decision Trees, Random Forest, and Gaussian Na¨ıve Bayes, along with deep learning models built using TensorFlow/Keras. To address class imbalances, the SMOTE technique is employed, while OpenCV handles image processing. Additionally, global variables are defined to store dataset paths, model parameters, and preprocessing configurations, ensuring smooth access to critical information throughout the application.

C. Key Functionalities

The system follows a structured pipeline for dataset handling and classification. The data upload and preprocessing module enables users to select datasets, extract image data, resize and normalize inputs, and apply class balancing techniques. Preprocessed datasets are split into training and testing sets to ensure model generalization. The model training module supports multiple classifiers, including Na¨ıve Bayes, Random Forest, and CNN-RFC hybrid models. It evaluates model performance using accuracy, precision, recall, and F1-score while generating confusion matrices for error analysis. The CNN model is trained with convolutional layers for feature extraction, followed by an RFC for final classification, ensuring robust anomaly detection.

To visualize results effectively, the system incorporates performance plotting functions, which generate accuracy and loss graphs during model training. The prediction module allows users to upload test images, process them through the CNN feature extractor, and classify them using the hybrid model. The results are displayed on the GUI with overlaid predictions. Additionally, a comparison module aggregates classifier performance metrics and visualizes them through bar charts, highlighting the superior accuracy of the hybrid CNN-RFC model over traditional methods.

D. Dataset Description

The dataset consists of image-based data organized into class-specific directories, representing different states such as fault and normal grid conditions. Each image is resized to a standard 64×64 pixel format to maintain consistency across the dataset. Preprocessing techniques include flattening, normalization, and SMOTE-based class balancing, ensuring that minority classes are adequately represented. The dataset is divided into training and testing subsets, with preprocessed arrays stored as NumPy files to facilitate quick loading during subsequent application runs.

E. Results and Analysis

The graphical user interface (GUI) provides an interactive platform for monitoring model performance at different stages. Upon dataset upload, the system confirms successful loading and displays class distributions. The preprocessing phase involves dataset balancing and transformation, with log outputs indicating the number of training and testing samples. The model training section presents detailed evaluation metrics, including accuracy scores, confusion matrices, and class distribution plots.

Experimental results demonstrate that the proposed CNN-RFC hybrid model significantly outperforms traditional classifiers. Compared to Na¨ıve Bayes and standalone Random Forest models, the hybrid approach achieves higher classification accuracy and better resilience against adversarial perturbations. Performance comparison charts illustrate incremental improvements in precision, recall, and F1-score across different models. Confusion matrices further validate the hybrid model's ability to distinguish between normal and fault conditions with minimal misclassifications.

The GUI also supports real-time classification, allowing users to upload test images and receive instant predictions. Sample outputs visually highlight the model's capability to detect grid anomalies accurately. The hybrid CNN-RFC model's effectiveness is evident in its ability to extract meaningful spatial features, classify complex patterns, and enhance cybersecurity in smart grids.

By integrating deep learning with ensemble learning, this research provides a robust solution for enhancing the security, accuracy, and adaptability of smart grid monitoring systems. The proposed approach offers a promising advancement in real-time anomaly detection, paving the way for more re- silient supervisory protection mechanisms in modern power networks. Table I presents comparative performance metrics.

TABLE I
PERFORMANCE COMPARISON OF MODELS

Model	Accuracy	Precision	Recall	F1-score
GNB	93.82%	95.09%	92.85%	93.57%
RFC	97.53%	98.11%	96.66%	97.31%
CNN-RFC	98.76%	99.03%	98.33%	98.67%

VI. RESULTS AND DISCUSSION

A. Implementation Description

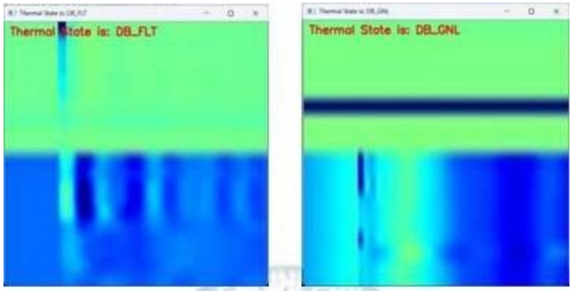
The proposed project is a comprehensive application that integrates data preprocessing, model training, performance evaluation, and real-time prediction functionalities through an intuitive graphical user interface (GUI). Built using Python's Tkinter library, the GUI allows users to interact seamlessly with the system while leveraging machine learning and deep learning libraries such as scikit-learn, imbalanced-learn, TensorFlow/Keras, and OpenCV for data processing and classification. The system comprises multiple modules, including data handling, model training, evaluation, and visualization, ensuring a streamlined workflow from dataset selection to final classification.

B. Results and Analysis

The graphical user interface (GUI) provides an interactive platform for monitoring model performance at different stages. Upon dataset upload, the system confirms successful loading and displays class distributions. The preprocessing phase involves dataset balancing and transformation, with log outputs indicating the number of training and testing samples.

The model training section presents detailed evaluation metrics, including accuracy scores, confusion matrices, and class distribution plots.

Experimental results demonstrate that the **proposed CNN-RFC hybrid model significantly outperforms traditional classifiers**. Compared to Na¨ıve Bayes and standalone Random Forest models, the hybrid approach achieves higher classification accuracy and better resilience against adversarial perturbations. Performance comparison charts illustrate incremental improvements in precision, recall, and F1-score across different models. Confusion matrices further validate the hybrid model's ability to distinguish between normal and fault conditions with minimal misclassification.


Thermal State is: DB_FLT

Thermal State is: DB_GNL

Fig. 3. Sample prediction outputs obtained using the proposed hybrid CNN with RFC model.

The GUI also supports real-time classification, allowing users to upload test images and receive instant predictions. Sample outputs visually highlight the model's capability to detect grid anomalies accurately. The **hybrid CNN-RFC model's effectiveness is evident in its ability to extract meaningful spatial features, classify complex patterns, and enhance cybersecurity in smart grids**.

By integrating deep learning with ensemble learning, this research provides a robust solution to improve the security, accuracy, and adaptability of smart grid monitoring systems. The proposed approach offers a promising advancement in real-time anomaly detection, paving the way for more re- silient supervisory protection mechanisms in modern power networks.

VII. CONCLUSION

This research proposed a hybrid model that integrates a Convolutional Neural Network (CNN) for robust feature extraction with a Random Forest Classifier (RFC) for decision-making. The model has demonstrated significant improvements in accuracy, precision, recall, and F1-scores when compared to traditional models such as Gaussian Na¨ıve Bayes (GNB) and standalone RFC classifiers in the context of supervisory protection systems for smart grids. This integrated approach effectively addresses the limitations inherent in conventional systems, including the inability to capture complex nonlinear patterns in Phasor Measurement Unit (PMU) data and the susceptibility to adversarial attacks. By leveraging

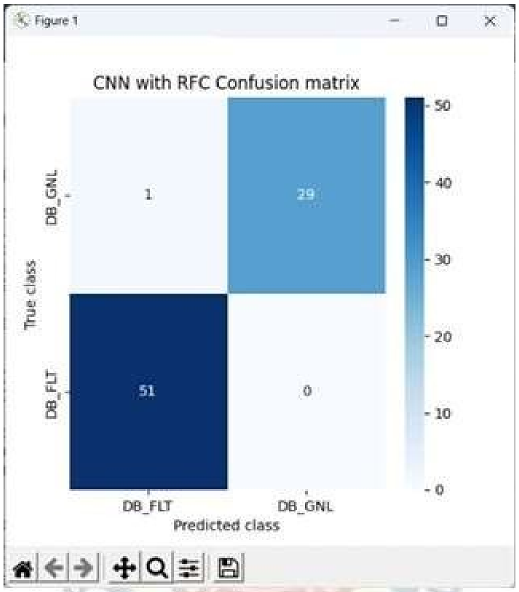


Fig. 4. Confusion Matrix obtained using the RFC model.

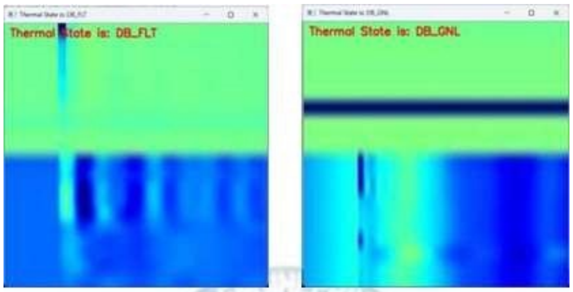


Fig. 5. Sample Output that the Model Predicted.

the CNN's ability to learn hierarchical feature representations from raw image data and combining it with the ensemble robustness of the RFC, the hybrid model achieves higher performance metrics and offers enhanced resilience against subtle data manipulations that can lead to misclassification and potential grid instability. The improved performance is clearly reflected in both the quantitative evaluations and the visual outputs, such as confusion matrices and performance comparison charts, which illustrate the hybrid model's superior capability in distinguishing between normal operations and attack scenarios.

VIII. Future Scope

Future research on adversarial attacks in deep learning-based supervisory protection systems for wide-area monitoring (PMU data) must embark on a revolutionary transformation, pioneering groundbreaking defense mechanisms that redefine cybersecurity in smart grids. The next frontier demands an unparalleled fusion of adversarial training, cutting-edge anomaly detection, and the unparalleled transparency of explainable AI to create impenetrable fortresses against sophisticated cyber threats. The integration of deep learning with physics-based models will not only enhance resilience but also forge an indestructible synergy between artificial intelligence and real-world grid dynamics.

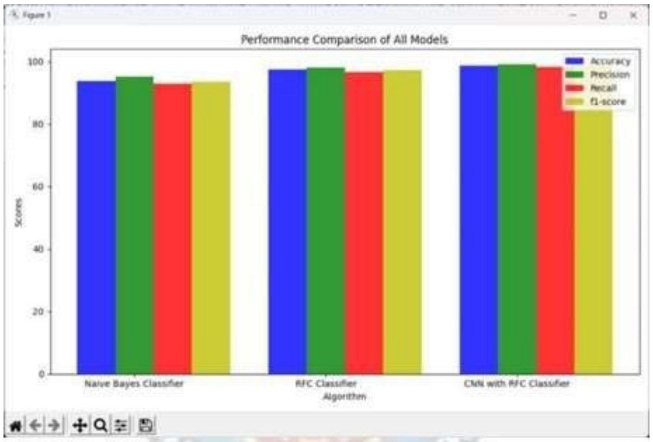


Fig. 6. Performance comparison of GNB classifier, RFC model, and proposed hybrid CNN with RFC model.

To combat ever-evolving attacks, real-time mitigation strategies must reach unprecedented heights, employing anomaly detection techniques so advanced that no intrusion goes unnoticed, alongside blockchain-powered, tamper-proof data transmission that sets a new gold standard for security. The exploration of adversarial attack transferability across diverse grid models will unveil hidden vulnerabilities, equipping the industry with predictive capabilities that anticipate and neutralize threats before they manifest.

Furthermore, the reinforcement of cyber-physical security through game-theoretic strategies and robust security frameworks will elevate grid protection to a masterclass of strategic defense. The ultimate breakthrough lies in the creation of self-learning AI models—hyper-intelligent systems that dynamically evolve, outpacing cyber adversaries with unparalleled adaptability and foresight. By harnessing the full potential of these innovations, the future of smart grids will not only be secured but will stand as an unassailable stronghold of technological supremacy, ensuring resilience, stability, and unmatched security for generations to come.

References

[1] Vahidi, S.; Ghafouri, M.; Au, M.; Kassouf, M.; Mohammadi, A.; Debbabi, M. Security of Wide-Area Monitoring, Protection, and Control (WAMPAC) Systems of the Smart Grid: A Survey on Challenges and Opportunities. IEEE Commun. Surv. Tutor. 2023, 25, 1294–1335. [2] Bu, S.; Meegahapola, L.G.; Wadduwage, D.P.; Foley, A.M. Stability and Dynamics of Active Distribution Networks (ADNs) with D-PMU Technology: A Review. IEEE Trans. Power Syst. 2022, 38, 2791–2804. [3] Berghout, T.; Benbouzid,

M.; Muyeen, S.M. Machine Learning for Cybersecurity in Smart Grids: A Comprehensive Review-Based Study on Methods, Solutions, and Prospects. Int. J. Crit. Infrastruct. Prot. 2022, 38, 100547. [4] Inayat, U.; Zia, M.F.; Mahmood, S.; Berghout, T.; Benbouzid, M. Cybersecurity Enhancement of Smart Grid: Attacks, Methods, and Prospects. Electronics 2022, 11, 3854. [5] Baba, M.; Nor, N.B.M.; Sheikh, A.; Nowakowski, G.; Masood, F.; Rehman, M.; Irfan, M.; Arefin, A.A.; Kumar, R.; Momin, B. A Review of the Importance of Synchrophasor Technology, Smart Grid, and Applications. Bull. Pol. Acad. Sci. Tech. Sci. 2022, 70, e143826. [6] Paramo, G.; Bretas, A.; Meyn, S. Research Trends and Applications of PMUs. Energies 2022, 15, 5329.

[7] Zhang, M.; Shen, C.; He, N.; Han, S.; Li, Q.; Wang, Q.; Guan, X. False Data Injection Attacks against Smart Gird State Estimation: Construction, Detection and Defense. Sci. China Technol. Sci. 2019, 62, 2077–2087. [8] Ravinder, M.; Kulkarni, V. A Review on Cyber Security and Anomaly Detection Perspectives of Smart Grid. In Proceedings of the 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 23–25 January 2023; pp. 692–697. [9] Lal, M.D.; Varadarajan, R. A Review of Machine Learning Approaches in Synchrophasor Technology. IEEE Access 2023, 11, 33520–33541. 38 [10]

Zhang, Y.; Shi, X.; Zhang, H.; Cao, Y.; Terzija, V. Review on Deep Learning Applications in Frequency Analysis and Control of Modern Power System. Int. J. Electr. Power Energy Syst. 2022, 136, 107744. [11] Bitirgen, K.; Filik, Ü.B. A Hybrid Deep Learning Model for Discrimination of Physical Disturbance and Cyber-Attack Detection in Smart Grid. Int. J. Crit. Infrastruct. Prot. 2023, 40, 100582. [12] Mississippi State University Critical Infrastructure Protection Center, Industrial Control System Cyber Attack Dataset. Available online: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data sets (accessed on 24 Jan 2025). [13] Chawla, A.; Agrawal, P.; Panigrahi, B.K.; Paul, K. Deep-Learning-Based Data Manipulation Attack Resilient Supervisory Backup Protection of Transmission Lines. Neural Comput. Appl. 2023, 35, 4835–4854. [14] Al-Hinai, A.S. Voltage Collapse Prediction for Interconnected Power Systems. Master's Thesis, West Virginia University, Morgantown, WV, USA, 2000. [15] Jahangir, H.; Lakshminarayana, S.; Maple, C.; Epiphaniou, G. A Deep Learning Based Solution for Securing the Power Grid against Load Altering Threats by IoT Enabled Devices. IEEE Internet Things J. 2023, 10, 10687–10697. [16] IEEE 14-Bus System. Available online: https://icseg.iti.illinois.edu/ieee-14-bus system/: :text=The (accessed on 25 Jan 2025). [17] Pai, A. Energy Function Analysis for Power System Stability; Springer Science Business Media: Berlin/Heidelberg, Germany, 1989. [18] Radhoush, S.; Vannoy, T.; Liyanage, K.; Whitaker, B.M.; Nehrir, H. Distribution System State Estimation and False Data Injection Attack Detection with a Multi Output Deep Neural Network. Energies 2023, 16, 2288. [19] Dolatabadi, S.H.; Ghorbanian, M.; Siano, P.; Hatziargyriou, N.D. An Enhanced IEEE 33 Bus Benchmark Test System for Distribution System Studies. IEEE Trans. Power Syst. 2021, 36, 2565–2572. [20] Lal, A. IEEE 69 Bus System. Available online: https://www.mathworks.com/matlabcentral/fileexchange/88111-ieee-69 bus-system (accessed on 25 Jan 2025). [21] Raghuvamsi, Y.; Teeparthi, K. Detection and Reconstruction of Measurements against False Data Injection and DoS Attacks in Distribution System State Estimation: A Deep Learning Approach. Measurement 2023, 210, 112565.