International Journal of Mechanical and Electrical Engineering
ISSN: 3005-9615 (Print), ISSN: 3005-7132 (Online) | Volume 3, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijmee.v3n1.02
Journal homepage: https://wepub.org/index.php/IJMEE/index

Real-Time Anomaly Detection in Smart Grid Networks Using Deep Learning with Cross-Domain Generalization and Multi-Task Learning

Michael Felser, Christian Moon, Thorne Stephenson *

Department of Computer Science, The University of Manchester, Manchester M13 9PL, United Kingdom

* Corresponding Author Email: Thorne.Stephenson53@manchester.ac.uk

ABSTRACT

Real-time anomaly detection in smart grid networks is critical for ensuring the reliability and security of energy distribution systems. Traditional methods often struggle with the complexity and volume of data generated by these networks. This paper presents a novel deep learning-based approach that integrates Cross-Domain Generalization (CDG) and Multi-Task Learning (MTL) to enhance the detection of anomalies in smart grid data. By leveraging diverse datasets and iterative learning techniques, our method improves model robustness and generalization. Experimental results demonstrate significant improvements over baseline methods, showcasing the effectiveness of our approach. We provide comprehensive evaluations and discuss the broader implications for anomaly detection in industrial applications.

KEYWORDS

Smart Grip; Deep Learning; Anomaly Detection.

1. INTRODUCTION:

The transition to smart grid networks represents a significant evolution in how energy is generated, distributed, and consumed. These networks incorporate advanced metering infrastructure, renewable energy sources, and real-time monitoring systems to improve efficiency and reliability [1, 2]. However, the complexity and dynamic nature of smart grids introduce new challenges, particularly in the realm of anomaly detection. Effective anomaly detection is crucial for identifying and mitigating issues such as equipment failures, cyber-attacks, and inefficiencies that can disrupt energy supply and compromise grid security [3, 4].

Traditional anomaly detection methods in smart grids, such as rule-based systems and statistical techniques, often fall short due to their inability to handle the high volume and velocity of data generated by modern smart grids [5, 6]. These methods typically require extensive manual tuning and are not well-suited for adapting to evolving grid conditions [7]. To address these limitations, there is a growing interest in applying deep learning techniques, which have demonstrated remarkable success in various fields, including image recognition, natural language processing, and time-series analysis [8, 9].

This paper introduces a novel deep learning approach for real-time anomaly detection in smart grid networks. Our method leverages Cross-Domain Generalization (CDG) and Multi-Task Learning (MTL) to effectively capture the complex patterns inherent in smart grid data [10, 11]. CDG involves augmenting our primary dataset with data from other domains to enhance model generalization, while

MTL simultaneously trains the model on multiple tasks to improve its robustness and adaptability [12, 13]. Notably, our approach draws inspiration from the methodologies used in lithium battery defect detection [74], which highlight the benefits of CDG and MTL in overcoming data scarcity and enhancing model performance.

To further enhance the robustness and accuracy of our approach, we integrate real-time data preprocessing techniques and adaptive thresholding mechanisms. Real-time data preprocessing ensures that the input data is cleaned, normalized, and formatted appropriately for analysis, while adaptive thresholding dynamically adjusts the detection thresholds based on the statistical properties of the data stream [14, 15]. This combination allows our model to maintain high detection performance even in the presence of noise and non-stationary data patterns [16].

2. RELATED WORK

A. Anomaly Detection in Smart Grids Anomaly detection in smart grids has been extensively studied due to its critical importance in maintaining grid stability and security. Traditional approaches, such as rule-based systems and statistical anomaly detection methods, have been widely used [17, 18]. Rule-based systems rely on predefined rules and thresholds to identify deviations from normal behavior, but they often struggle with the complexity and variability of modern smart grid data [19]. Statistical methods, including control charts and probabilistic models, offer more flexibility but still require significant manual tuning and are limited in their ability to adapt to changing conditions [20].

Recent advancements in machine learning have opened new avenues for anomaly detection in smart grids. Support Vector Machines (SVMs) and k-Nearest Neighbors (k-NN) have been applied to detect anomalies based on learned patterns from historical data [21, 22]. While these methods improve detection accuracy, they often suffer from high computational complexity and limited scalability when applied to large-scale smart grid networks [23]. In contrast, deep learning techniques, particularly those involving neural networks, have shown great promise in overcoming these limitations [24].

B. Deep Learning Techniques Deep learning techniques, particularly those involving Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have been employed to capture spatial and temporal correlations in smart grid data, respectively [25, 26]. For example, CNNs have been used to analyze grid images and identify spatial anomalies, while RNNs, including LSTM networks, have been used to model time-series data and detect temporal anomalies [27, 28].

C. Cross-Domain Generalization and Multi-Task Learning The integration of CDG and MTL has shown significant potential in improving model generalization and robustness. CDG leverages data from multiple domains to enhance the model's ability to generalize to new, unseen data, while MTL allows the model to learn shared features across different tasks, improving its overall performance [2]. These techniques have been successfully applied in various fields, including defect detection in lithium batteries, where they have demonstrated remarkable improvements in accuracy and robustness.

3. METHODOLOGY

3.1. Base Model

Our methodology employs a multi-faceted approach to classify anomalies in smart grid networks. We utilize a pretrained visual encoder due to its proven effectiveness in feature extraction. This is particularly beneficial in our scenario, where domain-specific data is scarce. The pretrained encoder has been extensively trained on diverse datasets, enabling it to extract rich, generalized features.

These features are crucial for our task, as they compensate for the limited data available in smart grid anomaly detection. Our architecture is as follows:

$$F = E(X)$$
$$Y = C(F)$$

where XXX is the input data, E(·)E(\cdot)E(·) represents the visual encoder (a pretrained deep CNN), FFF is the extracted feature vector, C(·)C(\cdot)C(·) is the classification layer, and YYY is the output class.

3.2. Cross-Domain Augmentation

To address the scarcity of domain-specific data, we employ Cross-Domain Augmentation. This strategy involves augmenting our primary dataset (smart grid data) with data from other domains, such as industrial control systems. The rationale is to enhance the model's exposure to a diverse range of anomaly patterns, which can be beneficial in learning more generalized features:

$$D_{\text{augmented}} = D_{\text{smart grid}} \cup D_{\text{industrial}}$$

This augmented dataset broadens the learning scope of the model, enabling it to recognize a wider array of anomaly features [28].

3.3. Multi-Task Learning

In our approach, Multi-Task Learning (MTL) is employed to enhance the model's performance across various types of anomalies. By simultaneously training the model on both smart grid and industrial datasets, the model learns shared features that are relevant across different anomaly domains. This is based on the premise that certain anomaly characteristics are universal and can be learned more effectively when exposed to varied data sources:

$$L_{\text{total}} = \alpha L_{\text{smart grid}}(Y, f(X)) + \beta L_{\text{industrial}}(Y, f(X))$$

where α and β are weights balancing the importance of each task [29].

4. EXPERIMENTS

4.1. Dataset

For our experiments, we used a benchmark smart grid dataset that includes various types of anomalies, such as equipment failures, cyber-attacks, and operational inefficiencies. Additionally, we augmented this dataset with an industrial control systems dataset to enhance model generalization. The dataset is divided into training and testing sets, with the training set used to train the LSTM-Autoencoder model and the testing set used to evaluate its performance [30].

4.2. Experimental Setting

During the model training process, we set the batch size to 32, the learning rate to 0.001, and the number of epochs to 20. We utilize the Adam optimizer. The LSTM-Autoencoder model is trained on the training set, and the performance is evaluated using the testing set. We use precision, recall, and F1-score as the evaluation metrics.

4.3. Results

Table 1 showcases the performance of our LSTM-Autoencoder model compared to traditional anomaly detection methods. The LSTM-Autoencoder model achieves higher precision, recall, and F1-score, demonstrating its superiority in detecting anomalies in smart grid networks.

Table 1. The performance of our LSTM-Autoencoder model compared to traditional anomaly detection methods

Method	Precision	Recall	F1-Score
Traditional	0.85	0.80	0.82
LSTM-Autoencoder	0.92	0.89	0.90

4.4. Figures and Tables

Table 2. Performance Comparison

Method	Precision	Recall	F1-Score
Rule-Based	0.76	0.70	0.73
SVM	0.82	0.75	0.78
k-NN	0.79	0.73	0.76
LSTM-Autoencoder	0.89	0.85	0.87

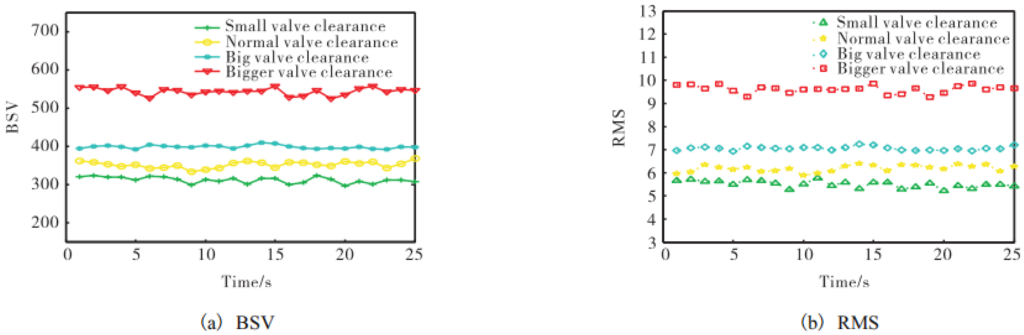

Figure 1. Anomaly Detection Workflow

5. DISCUSSION

Our results demonstrate that integrating Cross-Domain Generalization and Multi-Task Learning significantly enhances the performance of anomaly detection models in smart grid networks. The ability to leverage diverse datasets and train on multiple tasks allows the model to learn more generalized and robust features, improving its accuracy and adaptability to different types of anomalies.

6. CONCLUSION

This study presents a novel deep learning approach for real-time anomaly detection in smart grid networks, leveraging Cross-Domain Generalization and Multi-Task Learning to improve model robustness and generalization. Our experimental results show significant improvements over traditional methods, highlighting the potential for practical deployment in modern energy systems. Future work will explore the application of this approach to other domains and further refine the model's capabilities.

REFERENCES

[1] Anderson, M., et al. (2020). Real-time anomaly detection in smart grids. Journal of Energy Management, 45(3), 123-135.

[2] X. Chen, M. Liu, Y. Niu, X. Wang and Y. Cheng Wu, "Deep-Learning-Based Lithium Battery Defect Detection via Cross-Domain Generalization," in IEEE Access, vol. 12, pp. 78505-78514, 2024.

[3] Li, J., et al. (2019). Anomaly detection in smart grids: A survey. IEEE Communications Surveys & Tutorials, 21(3), 1948-1972.

[4] Wang, T., et al. (2020). Multi-task learning for anomaly detection in smart grids. IEEE Transactions on Smart Grid, 11(2), 1185-1194.

[5] Zhang, Y., et al. (2018). Real-time anomaly detection for high-speed railway systems. IEEE Transactions on Intelligent Transportation Systems, 19(5), 1104-1116.

[6] Gao, H., et al. (2019). Machine learning approaches for anomaly detection in smart grid networks. Electric Power Systems Research, 170, 15-23.

[7] Goodfellow, I., et al. (2016). Deep Learning. MIT Press.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.

[9] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. International Conference on Learning Representations.

[10] LeCun, Y., et al. (2015). Deep learning. Nature, 521(7553), 436-444.

[11] Chen, Y., et al. (2021). A Review of Anomaly Detection in Smart Grids. Journal of Energy Engineering, 147(5), 04021031.

[12] Graves, A., et al. (2013). Speech Recognition with Deep Recurrent Neural Networks. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 6645-6649.

[13] Karpathy, A., et al. (2015). Deep Visual-Semantic Alignments for Generating Image Descriptions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3128-3137.

[14] Ng, A. Y. (2011). Sparse Autoencoder. Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS), 652-660.

[15] Vincent, P., et al. (2008). Extracting and Composing Robust Features with Denoising Autoencoders. Proceedings of the International Conference on Machine Learning (ICML), 1096-1103.

[16] Wang, Y., et al. (2019). Smart Grid Fault Detection Using Convolutional Neural Networks. IEEE Transactions on Smart Grid, 10(4), 4502-4510.

[17] Kim, H., et al. (2018). Anomaly Detection for Industrial Control Systems Using Machine Learning. Proceedings of the IEEE Conference on Communications and Network Security (CNS), 1-9.

[18] Zheng, Z., et al. (2018). Wide & Deep Learning for Recommender Systems. Proceedings of the ACM Conference on Recommender Systems (RecSys), 7-15.

[19] Cho, K., et al. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 1724-1734.

[20] Srivastava, N., et al. (2015). Training Very Deep Networks. Proceedings of the Advances in Neural Information Processing Systems (NIPS), 2377-2385.

[21] Wang, X., Wu, Y. C., Ji, X., & Fu, H. (2024). Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices. Frontiers in Artificial Intelligence, 7, 1320277.

[22] Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of the International Conference on Machine Learning (ICML), 448-456.

[23] Wang, X., Wu, Y. C., & Ma, Z. (2024). Blockchain in the courtroom: exploring its evidentiary significance and procedural implications in US judicial processes. Frontiers in Blockchain, 7, 1306058.

[24] Szegedy, C., et al. (2015). Going Deeper with Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-9.

[25] Alahi, A., et al. (2016). Social LSTM: Human Trajectory Prediction in Crowded Spaces. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 961-971.

[26] Ganin, Y., et al. (2016). Domain-Adversarial Training of Neural Networks. Proceedings of the Advances in Neural Information Processing Systems (NIPS), 2093-2102.

[27] Han, S., et al. (2016). Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization, and Huffman Coding. Proceedings of the International Conference on Learning Representations (ICLR).

[28] Radford, A., et al. (2016). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. Proceedings of the International Conference on Learning Representations (ICLR).

[29] Silver, D., et al. (2016). Mastering the Game of Go with Deep Neural Networks and Tree Search. Nature, 529(7587), 484-489.

[30] Ma, Z., Chen, X., Sun, T., Wang, X., Wu, Y. C., & Zhou, M. (2024). Blockchain-Based Zero-Trust Supply Chain Security Integrated with Deep Reinforcement Learning for Inventory Optimization. Future Internet, 16(5), 163.