IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 December 2024, accepted 27 December 2024, date of publication 30 December 2024, date of current version 6 January 2025.

Digital Object Identifier 10.1109/ACCESS.2024.3524061

RESEARCH ARTICLE

Enhancing Power Grid Reliability With Machine Learning and Auxiliary Classifier Generative Adversarial Networks: A Study on Fault Detection Using the Georgia Electric System Load Dataset

HAFEEZ UR REHMAN SIDDIQUI [1], (Student Member, IEEE), ROBERT BROWN [2], ADIL ALI SALEEM [1], MUHAMMAD AMJAD RAZA [1,3], AND SANDRA DUDLEY [2], (Member, IEEE)

[1]Institute of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, Punjab 64200, Pakistan
[2]Bioengineering Research Centre, School of Engineering, London South Bank University, SE1 0AA London, U.K.
[3]Faculty or Computing, Riphah International University, Sahiwal 57000, Pakistan

Corresponding author: Hafeez Ur Rehman Siddiqui (Hafeez@kfueit.edu.pk)

This work was supported by the UKRI/EPSRC Network Plus "A Green Connected and Prosperous Britain," under Reference EP/W034204/1.

ABSTRACT Power networks are vital to society, yet service outages and faults can have devastating consequences. This study introduces a novel integration of machine learning and data augmentation techniques for fault detection and classification, addressing gaps in data diversity and imbalance. Unlike traditional approaches, the research utilizes an Auxiliary Classifier Generative Adversarial Network (ACGAN) to generate synthetic data representative of underrepresented fault types, enhancing model training and performance. By extracting both spectral and statistical features from the Grid Event Signature Library (GESL) dataset, a comprehensive representation of power system signals is achieved. A comparative evaluation of models including Decision Trees (DT), Random Forest (RF), Extra Tree Classifier (ETC), Gradient Boosting Classifier (GBC), and K-Nearest Neighbors (KNN) revealed the Extra Tree Classifier achieved the highest testing accuracy of 93.85%. The methodologies demonstrated scalability by using a dataset augmented to 9,000 samples and validated robustness through 10-fold cross-validation with a standard deviation of 0.00659. These results highlight the proposed framework's potential for real-world implementation in modern power grids, offering enhanced fault prediction and resilience. This research establishes a pathway for integrating advanced data augmentation and machine learning techniques into operational power grid systems, ensuring stability and reliability.

INDEX TERMS ACGAN, grid event signature library, smart grid, fault tolerance, Machine Learning, spectral features.

I. INTRODUCTION

A reliable power system is essential for modern society as almost every device, system, and service depends directly or indirectly on the electrical grid. The consequences of faults and service outages can be severe, including fires, explosions, intermittent disconnection of residents, reliance on backup power supplies for vital medical equipment, communication and transportation service outages, water pollution, and other wide reaching concerns. These issues may become more frequent as we move away from fossil-based energy sources to renewables. Therefore, it is crucial to incorporate sophisticated metering, real-time monitoring, and warning systems into the electricity network to enable an effective and resilient smart grid [1], [2]. Modern power grids can be intricate, ever-changing systems, consistently expanding to support ever changing industries, sectors and urban planning. Modern grids are developing via the incorporation of sustainable energy sources, intelligent grid technology, and sophisticated monitoring tools. The increasing complexity of power grid systems, caused by the increased use and need of distributed energy sources,

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

worsens temporary disruptions and exposes deficiencies in safety measures. Identifying, categorizing, and pinpointing sudden changes are crucial for enhancing the overall efficiency of a system and reducing disruptions in smart grid environments [3], [4]. High impedance problems, such as ''arcing'', cause significant difficulties. Power line problems in recent years have caused devastating wildfires, including Dixie, Thomas, Camp, Red-wood Valley, Atlas, and Nuns in the United States alone. These fires have resulted in considerable loss of life and extensive economic damage [5]. While fires caused by ''equipment failure'' or ''electrical power'' are approximately 20% of events, their environmental consequences are disproportionately severe, resulting in the combustion of nearly 1.5 million acres according to the above study [5]. To ensure the stability and dependability of power systems, advanced methods for detecting, predicting, and categorizing faults are necessary to ensure proactive management.

The Grid Event Signature Library (GESL) [6] is a major resource in this field, including a comprehensive collection of real-world labelled fault data. It supports the development and evaluation of innovative approaches to improve power system resilience. Prediction and categorization of faults are crucial elements of power system management. Operators can successfully minimize downtime and mitigate the danger of cascade failures by accurately predicting probable faults and implementing preventative actions. Fault categorization facilitates a prompt and suitable response to abnormalities, guaranteeing that repair measures are customized to the particular nature of the disruption and the scene. The dataset provided by GESL includes a diverse collection of power system events, ranging from 'simple disruptions like voltage sags and swells to more intricate phenomena such as oscillations and transients' [6]. Disruptions in electrical power networks can range in severity from minor to severe. Voltage sags refer to temporary decreases in voltage, usually induced by the initiation of large motors or electrical problems. On the other hand, voltage swells are temporary rises in voltage that occur when a load is suddenly disconnected. Both have the potential to cause harm to delicate machinery. Oscillations refer to periodic fluctuations in voltage or current indicate possible instability in the system, while transients, which are brief, high-frequency surges are induced by occurrences such as lightning strikes or switching processes. Transient electrical disturbances can result in substantial harm to electrical equipment, highlighting the importance of implementing resilient power system management. The variety of GESL makes it well-suited for training and testing machine learning models that focus on fault prediction and classification.

A. CONTRIBUTIONS

This study makes the following key contributions:

- Comprehensive extraction and integration of power system event data from the GESL, ensuring a robust dataset for analysis. The dataset includes a wide range of power system occurrences and sensor data, providing a solid foundation for developing and evaluating machine learning models.
- Extraction of a diverse set of spectral and statistical features from the sensor data, enhancing the models' ability to capture both time-domain and frequency-domain features of the power system signals.
- Utilisation of the ACGAN to augment the GESL dataset, significantly increasing the number of samples and improving data diversity and representativeness. This augmentation addresses the issue of data imbalance by generating high-quality synthetic data for underrepresented classes.
- Application and evaluation of several machine learning models, including Decision Trees (DT), Random Forest (RF), Extra Tree Classifier (ETC), Gradient Boosting Classifier (GBC), and K-Nearest Neighbors (KNN), on both the original and augmented datasets. The evaluation metrics include training accuracy, testing accuracy, precision, recall, and F1-score.

II. LITERATURE REVIEW

The GESL dataset has become a vital resource for power system research, providing comprehensive real-event data, essential for developing and validating detection methods. Studies utilizing GESL data have highlighted its adaptability and significance in both academic and industrial contexts. For instance, [6] introduced a novel event signature library using a Deep Neural Network (DNN) to classify electrical disturbances, enhancing traditional event classifications. This method demonstrated a significant capability to differentiate between event types, with over 1,000 occurrences analyzed. In [7], a big data framework was developed to categorize power grid signatures, improving intelligent learning through AI and ML approaches. This library now contains over 800 signatures and 250 million data points, supporting advanced analysis and forecasting for grid reliability. The Signature Matching Tool (SMT) described in [8] aids users in recognizing electrical signatures by employing a Local Binary Classifier (LCN), achieving an average accuracy of 83% and reducing training time compared to traditional classifiers. Meanwhile, [9] presented the Spectral Correlation Function (SCF) for identifying grid-signal distortions, outperforming fast fourier transform (FFT) and power spectral density (PSD) methods in detecting signal types. Researchers in [10] introduced a method to define thresholds for oscillation detection, enhancing automation and reliability in identifying significant oscillations. This approach minimizes false alarms while maintaining precision. The authors in [5] proposed an energy detection algorithm for smart grids, improving disruption identification and automatic warnings using real data from GESL and the DOE/EPRI National Database. The performance of various sensors in power systems was evaluated in [11], revealing that advanced sensors may excel in detecting high-frequency transients compared to traditional current transformers. The same researchers highlighted an energy detector algorithm's efficiency in identifying transient events with high accuracy [12]. MENSA, an Intrusion Detection System (IDS) discussed in [13], combines Autoencoder and GAN architectures to detect and classify cyber intrusions in smart grids with high accuracy. Lastly, [14] introduced an extensive open-source dataset of synchrophasor measurements, offering valuable data for benchmarking and developing new algorithms.

Generative adversarial networks (GANs) have demonstrated significant potential in addressing data scarcity and imbalance in energy fault prediction and related fields. Recurrent GANs (R-GANs), which use Recurrent Neural Network (RNNs) to capture temporal dependencies, were applied to the UCI Appliances Energy Prediction and Building Data Genome datasets, achieving MAPE values of 10.12% (TSTR) and 10.81% (TRTR) for the UCI dataset, demonstrating comparable performance between synthetic and real data models [15]. Auxiliary Classifier GANs (ACGANs) generated realistic labeled sensor data for mechanical fault diagnosis, achieving 100% fault classification accuracy on vibration signal data, even with unbalanced datasets [16]. Enhanced GANs (E-GANs) combined DCGANs with Convolutional Neural Networks (CNNs), incorporating k-means clustering and ridge regression for feature optimization. On the gearbox dataset, E-GAN achieved 73.09% accuracy under a 1:20 imbalance ratio, outperforming DCGAN-based models. Conditional DCGANs (C-DCGANs) improved classification accuracy to 99.9% on the Case Western Reserve (CWRU) and planetary gearbox datasets by balancing datasets with synthetic samples [17]. In power grids, semi-supervised GAN frameworks like the generative-adversarial based semi-supervised learning framework (GBSS) integrated CGANs with Deep Ladder Networks, achieving high F-measure scores and robust performance in diagnosing faults and cyberattacks across imbalanced datasets with varying label ratios [18]. These studies illustrate GANs' ability to enhance fault prediction and classification in energy systems, particularly when historical data is scarce or imbalanced.

Overall, the literature highlights GESL's critical role in advancing power systems research. By leveraging this rich dataset, various sophisticated methods have been developed to enhance fault detection, prediction, and classification, demonstrating GESL's importance in improving the reliability and stability of power systems.

III. METHODOLOGY

This section provides a detailed explanation of the thorough process of gathering data and extracting significant features, which are essential to our research technique.

A. PROPOSED METHODOLOGY

The proposed methodology used in this manuscript involves several key steps: data extraction, preprocessing, feature extraction, data augmentation, and model training and evaluation. Data was extracted from the GESL using a web scraping tool and saved in CSV format. The dataset was then preprocessed to handle missing values and to standardize the data. Spectral and statistical features were extracted from the sensor data to capture both time-domain and frequency-domain characteristics. To address data imbalances, ACGAN was used for data augmentation, significantly increasing the dataset size. Machine learning models, including Decision Trees (DT), Random Forest (RF), Extra Tree Classifier (ETC), Gradient Boosting Classifier (GBC), and K-Nearest Neighbors (KNN), were trained and evaluated on both the original and augmented datasets. Model performance was assessed using metrics such as accuracy, precision, recall, F1-score, and cross-validation to ensure robustness and reliability.

B. DATA SET

The data used in this study was carefully extracted from the GESL website (https://gesl.ornl.gov/), which offers an all-encompassing interface for signatures. The dataset comprises a wide array of power system occurrences in 1D wave form, such as transients, voltage sags, surges, and oscillations. In the dataset, an individual signature ID is assigned to each record, which is accompanied by comprehensive metadata such as sources of data and event identifiers. A further subdivision of these main headings (Condition, Equipment, Events, Phase, and State) results in a total of 1485 unique event identifiers. By analyzing these event identifiers, this research seeks to discern patterns and gain insights regarding power system disturbances.

Aside from event tags, the collection also contains sensor data from multiple equipment. These sensors are essential for monitoring and evaluating power system events and consist of:

1. Frequency Disturbance Recorder (FDR): An FDR is a device that monitors and records the frequency variations in the power grid. Its primary role is to detect and analyze disturbances in the grid frequency, which can indicate issues such as load imbalances, generation losses, or system faults. By tracking these disturbances, FDRs help in maintaining the stability and reliability of the power grid.
2. Acoustic Emission on Side Wall: Acoustic emission refers to the release of transient elastic waves produced by a sudden redistribution of stress in a material. In the context of power systems, acoustic emission sensors on the side walls of equipment like transformers or generators can detect early signs of mechanical failures, partial discharges, or other anomalies. This helps in preventive maintenance and avoiding catastrophic failures.
3. Quadripole: In electrical engineering, a quadripole (or four-terminal network) is a two-port network with two pairs of terminals. It is used in the analysis of electrical circuits to describe input and output relationships. In power systems, quadripoles can be used to model and analyze the behavior of transmission lines and network components, facilitating the design and stability analysis of the power grid.

4. Ultra-High Frequency (UHF): UHF refers to the frequency range between 300 MHz and 3 GHz. In power systems, UHF sensors are often used to detect partial discharges in equipment like transformers, gas-insulated switchgear, and cables. Partial discharges are early indicators of insulation deterioration, and UHF detection allows for early intervention to prevent failures.
5. Acoustic Emission on Top Lid: Similar to acoustic emission on the side wall, sensors placed on the top lid of power system equipment monitor for stress-induced acoustic waves. These sensors help in detecting and locating faults, mechanical stresses, or partial discharges within the equipment, ensuring timely maintenance and reducing the risk of breakdowns.
6. Optical Sensors (OS): OSs are used for various power system applications, including temperature, current and voltage sensing, and detecting partial discharges. They provide high accuracy and immunity to electromagnetic interference, making them ideal for monitoring high-voltage equipment and ensuring operational safety and reliability.
7. High-Frequency Current Transformer (HFCT): HFCTs are used to detect high-frequency components of electrical currents, often associated with partial discharges or high-frequency transients in power systems. By monitoring these high-frequency signals, HFCTs help in identifying and diagnosing insulation issues and ensuring the health of critical power equipment.
8. Acoustic Emission: In general, acoustic emission in power systems refers to the monitoring of elastic waves generated by stress, deformation, or discharges within electrical equipment. This non-destructive testing method helps in early fault detection, condition monitoring, and preventive maintenance of assets like transformers, circuit breakers, and generators.
9. Potential Transformer/Current Transformer (PT/CT): PTs (Potential Transformers) and CTs (Current Transformers) are used for voltage and current measurement in power systems, respectively. PTs step down high voltages to safer levels for metering and protection, while CTs reduce high currents to measurable levels. Both are essential for accurate measurement, control, and protection of the power grid.
10. Phasor Measurement Unit (PMU): A PMU is a device that measures the electrical waves on an electricity grid to determine the magnitude and phase angle of the sinusoidal waveforms. PMUs provide real-time monitoring of the grid's voltage and current phasors, which are critical for grid stability analysis, fault detection, and dynamic system response assessment. PMUs play a key role in enhancing situational awareness and operational efficiency of modern power systems.

These components and technologies are integral to the monitoring, analysis, and maintenance of power systems, contributing to the overall reliability and efficiency of the electricity grid. Every sensor records different aspects of power system functionality, hence enhancing the overall comprehension of the factors that contribute to power outages. In cases where sensor data was missing, a 'NaN' label was used to denote the absence of data. The data was acquired through the use of a web scraping tool and subsequently saved in a CSV file titled "DataLabels.csv." In addition, sensor data can be accessed via an Application Programming Interface (API) and is stored in separate CSV files that are titled after their corresponding signature IDs (e.g., sigID-5.csv for signature ID 5). The format of these CSV files varies due to the diverse type, quantity of sensors, and length of the wave data. In order to maintain uniformity across the dataset, only one channel (voltage) was chosen for analysis from all the sensor data.

C. FEATURE EXTRACTION

The data from each sensor was carefully analyzed to extract an array of spectral and statistical features. This was carried out to improve the performance of machine learning (ML) models and enable more accurate analysis. The features that have been retrieved are Mean, Standard Deviation (StdDev), Skewness, Kurtosis, Maximum Power Spectral Density Frequency (Max PSD Freq), Spectral Centroid, Spectral Flux, and Spectral Decrease. The retrieved features are essential for capturing the properties of signals in both the time-domain and frequency-domain, providing a holistic perspective. Time-domain features, such as Mean and StdDev, offer a clear and easily understandable insight into the data, which is essential for comprehending model behavior and results. The Mean provides a measure of central tendency, serving as a reference point for the variation in the data, while StdDev quantifies the extent of variability, aiding in the identification of the stability or volatility within the signal. Frequency-domain features, such as Max PSD Freq and Spectral Centroid, are essential for comprehending the periodicity and frequency distribution of a signal. Max PSD Freq identifies the most prominent frequency component, which is very valuable for tasks like predictive maintenance and defect detection, where certain frequency patterns can indicate potential issues. The Spectral Centroid, which indicates the central point of the spectrum, offers valuable information about the distribution of high and low frequencies, facilitating the distinction between different types of signals.

Features such as Spectral Flux and Skewness play a crucial role in coping with the dynamic and unpredictable features found in time series data. Spectral Flux quantifies the speed of variation in the spectrum, allowing models to identify dynamic changes and transitions in the signal. It is crucial for recording fleeting events or sudden shifts in the signal's behavior. Skewness quantifies the lack of symmetry in the distribution of data, revealing any deviations and unexpected changes that are important for identifying anomalies. In addition, the inclusion of features such as Kurtosis and Spectral Decrease enhances the depth of the study. Kurtosis quantifies the degree of deviation from a normal distribution, providing insights into the existence of outliers, aiding models in predicting exceptional results. Spectral Decrease, a measure of the reduction in higher frequencies with time, provides valuable information about the harmonic composition and energy distribution of a signal. It helps to analyze the signal's timbre and detect any deterioration or damping effects. These features work together to accurately capture both the time-based and frequency-based features of the signals. This comprehensive approach improves the ability of ML models to distinguish between different patterns or classes. The extensive range of features enables models to successfully differentiate between various signal types or states, adjust to variations in data distribution over time, and deliver precise and reliable predictions.

D. DATA INTEGRATION AND FINAL DATASET

The feature extraction process incorporated labels obtained from the 'DataLabels.csv' file, using a common identifier column, 'sigID'. This dataset originally contained 50 distinct labels, each representing a different type of event. As depicted in Figure 1, the data distribution of these labels was highly uneven. The label with the highest number of instances appeared 241 times, while the label with the fewest occurrences appeared only once.

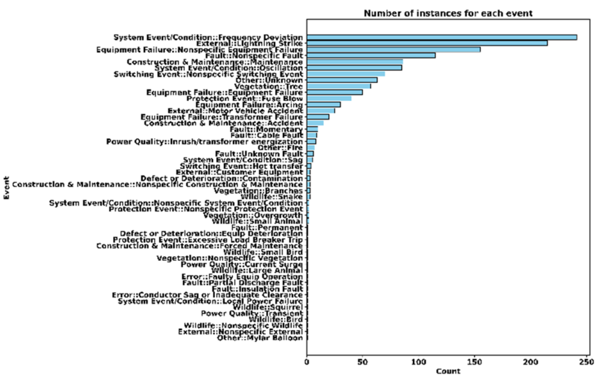


FIGURE 1. Original label distribution of the dataset.

An uneven distribution of data presents substantial difficulties for data interpretation and the training of models. Models trained on this dataset could exhibit a bias towards the more prevalent labels, perhaps leading to the neglect or misclassification of the less common ones.

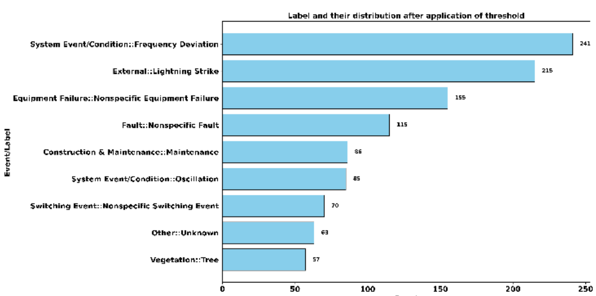


FIGURE 2. Labels and their distribution after application of threshold.

To tackle this problem, the authors implemented a criterion to exclude labels with fewer than 50 occurrences. The purpose of this choice was to establish a dataset that is both well-balanced and robust, ensuring that each label included has an adequate number of occurrences for dependable analysis and model training. By using this criterion, the dataset underwent a substantial reduction, prioritizing the most dominant labels. The dataset has been modified to contain only events that have 50 or more instances, as seen in Figure 2. This modification led to a more equitable allocation of categories, enabling more precise and dependable analysis of data and training of models.

After performing label encoding, instances with null values were excluded from the dataset, along with the 'sigID' column. The cleaned dataset, now devoid of null values and unnecessary identifiers, provides a more streamlined and analysis-ready version for subsequent processing. This preprocessed dataset, and snippet of which is shown in Figure 3, ensures that each entry is complete and labeled appropriately, facilitating more effective data analysis and machine learning model training.

Mean	Std Dev	Skewness	Kurtosis	Max PSD Freq	centroid	flux	decrease	Label
0.609588	0.330459	0.447303	0.000211	0.008938	0.014134	0.095047	0.467065	8
0.617707	0.000011	0.407018	0.044636	0.009314	0.319465	-0.000232	0.452362	6
0.620487	-0.000106	0.261243	0.242309	0.007957	0.383624	0.000002	0.421424	6
0.617475	0.000660	0.444796	0.083483	0.100050	0.476205	-0.000950	0.541896	3
0.614192	0.000057	0.364182	0.068570	0.017199	0.298620	-0.000183	0.454182	6
...	...	...	...	...	...	...	...	...
0.614890	0.000294	0.489585	0.363245	0.229666	0.609965	-0.000092	0.450033	5
0.619510	0.000139	0.497920	0.209490	0.626415	0.848554	-0.001212	0.348540	3
0.621316	-0.000026	0.460793	0.002651	0.012360	0.374160	-0.000341	0.355593	7
0.626403	-0.000064	0.475996	0.111294	0.021357	0.206331	-0.000016	0.417020	6
0.614103	0.002047	0.463709	0.196482	0.292199	0.657549	-0.003137	0.235163	1

FIGURE 3. Snippets of the dataset after label encoding and removal of null values.

The process of label encoding converts categorical labels into numerical values, which is essential for ML algorithms that require numerical input. This transformation helps in maintaining the intrinsic relationships between different labels, enabling the models to learn and predict more accurately. By removing instances with null values, the dataset's integrity and quality are preserved, minimizing the potential for erroneous data to affect analysis outcomes. Furthermore, the exclusion of the 'sigID' column, which served primarily as an identifier, removes unnecessary noise from the dataset. This step is crucial, as it prevents the inclusion of irrelevant data that could potentially skew results. The resulting dataset comprises 1087 samples and 9 features.

IV. DATA AUGMENTATION

The dataset used in this study, comprising 1087 samples and 9 features, presented significant challenges for effectively training the employed ML models. The limited size of the dataset was insufficient for the models to generalize efficiently, resulting in issues of overfitting and underfitting. The discrepancy with the models trained on such a dataset is discussed in the result and discussion section.

A significant effort was made to handle bias, particularly the class imbalance present in the dataset. Since the original dataset exhibited a highly uneven class distribution, the models are biased toward the majority class as shown in result and discussion section. To mitigate this, the ACGAN [19], [20], [21] was employed for data augmentation. ACGAN was chosen to artificially expand the dataset, generating more varied and representative samples that better captured the underlying data distribution. This augmentation process aimed to mitigate overfitting by providing a richer and more diverse dataset, enhancing the models' ability to generalise to new, unseen data. An ACGAN is a type of Generative Adversarial Network (GAN) [21], [22], [23] that is particularly effective for data augmentation tasks. In an ACGAN, the generator not only produces realistic data samples but also generates them conditioned on class labels, which allows for more controlled data generation and can enhance the diversity and quality of the augmented data. The ACGAN offers several notable advantages over the traditional GAN. One of the primary benefits of ACGAN is its ability to generate class-conditional samples. By incorporating class labels into the training process, ACGAN enables the generator to produce data samples that are not only realistic but also representative of specific classes. This controlled data generation capability is particularly useful in scenarios where the dataset is imbalanced, as it allows for the targeted augmentation of underrepresented classes, thereby improving the performance and robustness of machine learning models trained on the augmented data. Furthermore, the inclusion of an auxiliary classifier within the discriminator enhances the learning process. This classifier predicts the class labels of the input data, providing additional feedback to the generator. As a result, the generator receives more informative gradients, which helps it produce higher quality and more diverse samples. This dual objective (distinguishing real from fake samples and classifying the input data) makes the discriminator a more powerful and effective component in the network.

Additionally, ACGAN improves the overall stability and convergence of the training process. Traditional GANs often suffer from issues such as mode collapse, where the generator produces limited varieties of samples [24], [25]. ACGAN mitigates this problem by encouraging the generation of samples across different classes, thereby promoting diversity and reducing the likelihood of mode collapse, further demonstrating ACGAN as a reliable and robust approach for generating synthetic data, particularly in complex and multi-class scenarios. The architecture of the generator in the ACGAN used in this manuscript is presented in Figure 4 (a). This is a multi-layered neural network designed to transform input noise vectors into realistic data samples. This process is facilitated through a series of dense layers, each followed by batch normalisation and dropout layers to ensure effective learning and to prevent overfitting.

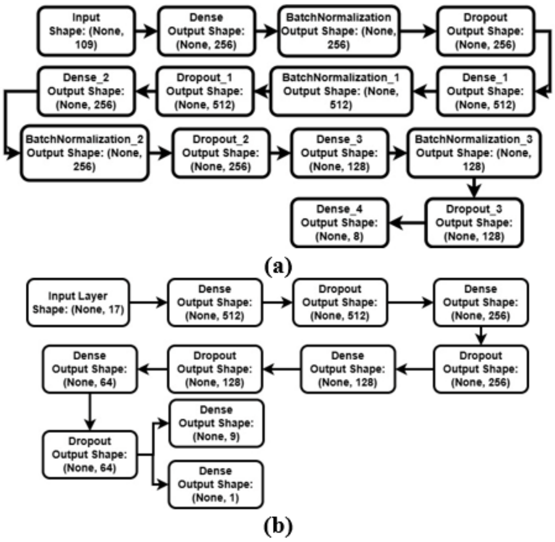


FIGURE 4. **Architecture of GAN used in this study (a) Discriminator (b) Generator.**

The generator starts with an input layer that accepts noise vectors of shape (None, 109). These noise vectors are the seeds from which the generator creates data samples. The first dense layer then processes these input vectors, expanding them to 256-dimensional feature representations. This expansion is crucial as it allows the network to learn complex transformations required for generating realistic samples. Following the first dense layer, batch normalization is applied. Batch normalization helps in stabilizing and accelerating the training process by normalizing the outputs of the dense layer, thus ensuring that the activations remain within a suitable range. This is immediately followed by a dropout layer, which randomly drops 50% of the units in the previous layer during training. Dropout is a regularization technique that helps prevent overfitting by ensuring that the network does not become overly reliant on any particular perceptrons.

The second dense layer further increases the feature dimensionality to 512, enabling the generator to capture more detailed and complex patterns. This layer is also followed by batch normalisation and dropout layers, maintaining the same pattern of stabilisation and regularisation seen in the previous stages. Subsequent to this, the third dense layer reduces the feature dimensionality back to 256, which is again normalised and regularised through batch normalisation and dropout. The alternation between expanding and contracting the feature space helps in refining the learned representations, allowing the network to generate more accurate and detailed samples. In the fourth stage, another dense layer is introduced, this time reducing the feature dimensionality to 128. This progressive reduction helps in converging the high-dimensional representations into more concise forms that are easier to manage and transform into the final output. Finally, the last dense layer transforms the 128-dimensional features into an 8-dimensional output. This final transformation step is crucial, as it shapes the output into the desired format, ready for further processing or evaluation.

The discriminator architecture of ACGAN used in this manuscript, presented in Figure 4(b), is also a sophisticated multi-layered neural network designed to differentiate between real and generated data samples while also classifying the input data into its respective classes. This dual functionality is achieved through a sequence of dense layers, each followed by dropout layers to ensure robust learning and mitigate overfitting. The discriminator begins with an input layer that accepts data samples of shape (None, 17). This input layer serves as the starting point for the network to process the features of the data samples. The first dense layer expands these input features to 512 dimensions, significantly increasing the capacity of the network to capture intricate patterns in the data. This layer is followed by a dropout layer, which drops 50% of the units randomly during training to prevent overfitting and enhance the network's generalization capabilities. The second dense layer continues the feature transformation, maintaining the 512-dimensional space. This consistency helps the network stabilize and build deeper feature representations. This layer is also followed by another dropout layer, reinforcing the regularization effect.

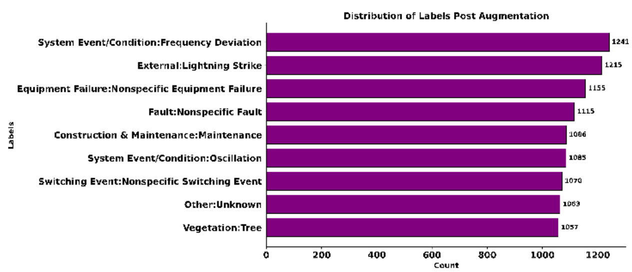


FIGURE 5. **Class label distribution post augmentation.**

The third dense layer reduces the feature dimensionality to 256, which allows the network to begin converging the high-dimensional features into more concise forms. Subsequently, a dropout layer is applied to ensure that the reduced dimensionality does not lead to overfitting. In the fourth stage, the network further reduces the feature dimensionality to 128 through another dense layer. This gradual reduction helps in refining the learned representations, making them more abstract and closer to the final output. This is followed by another dropout layer to maintain regularization. The fifth dense layer further reduces the dimensionality to 64, which is critical for simplifying the features into a form that can be easily classified. A subsequent dropout layer continues to enforce the regularization. Finally, the architecture branches into two separate dense layers to fulfill its dual objectives. The first branch consists of a dense layer that outputs a single value, representing the validity score (real or fake) of the input sample. The second branch includes a dense layer that outputs a vector of 9 dimensions, corresponding to the class probabilities of the input sample.

The use of ACGANs for data augmentation was a pivotal step in addressing the limitations of the original dataset. By generating high-quality synthetic data and merging it with existing data, the research achieved a more balanced and extensive dataset. The distribution of labels post-augmentation, as illustrated in Figure 5, demonstrates a more balanced representation of the various event types, crucial for ensuring the generalisation and effectiveness of the machine learning models trained on this data.

V. RESULTS AND DISCUSSION

A. CLASSIFICATION RESULTS ON THE ORIGINAL DATASET

The dataset used in this study comprises 1087 samples and 9 features, split into training and testing sets with an 80:20 ratio. Given the dataset's moderate size and the complexity inherent in its diverse sensor data, ensemble models such as Random Forest (RF), Extra Tree Classifier (ETC) and Gradient Boosting Classifier (GBC) were employed to ensure optimal predictive accuracy and robustness. The complexity of the dataset arises from its varied sensor data, with features encompassing both time-domain and frequency-domain. This necessitates robust preprocessing and the application of model's adept at managing diverse data characteristics. Ensemble methods like RF and Gradient Boosting strike an appropriate balance between model complexity and overfitting risk, offering high performance without the computational demands typically associated with deep neural networks. RF and ETC models are particularly effective in mitigating overfitting by averaging multiple decision trees (DT), elucidating feature importance, and handling non-linear relationships within the data. On the other hand, GBC models sequentially correct errors from preceding models, capturing subtle patterns and managing outliers adeptly. Although DT models are individually less powerful, they provide valuable interpretability and have low computational cost, making them suitable for baseline analysis. The study also evaluated the K-Nearest Neighbors (KNN) classifier, which, despite its simplicity and non-parametric nature, struggles with higher-dimensional data and noise typically present in such real-world datasets. The hyperparameters for each model, selected using grid search, are detailed in Table 1.

The models were trained on the training data and then tested on the testing set. The results of these evaluations are presented in Table 2.

The results presented in Table 2 from training and testing reveal several key insights. The high training accuracies of the DT and RF models (both at 0.979) contrast sharply with their significantly lower testing accuracies (0.445 and 0.486, respectively). This discrepancy suggests that these models are overfitting the training data, capturing noise and specific patterns that do not generalise well to unseen data. Similarly, the GBC model, with a training accuracy of 0.922 and a testing accuracy of 0.467, also indicates overfitting, although

TABLE 1. Hyperparameters used in this study.

Classifier	Hyperparameters
ETC	n_estimators: 100, random_state: 123
DT	criterion: 'gini', random_state: 123, splitter: 'best'
RF	bootstrap: True, criterion: 'gini', max_features: 'sqrt', n_estimators: 100, random_state: 123
GBC	criterion: 'friedman_mse', learning_rate: 0.1, loss: 'log_loss', max_depth: 3, n_estimators: 100, random_state: 123
KNN	n_neighbors: 5

TABLE 2. Results of classifiers on original dataset.

Model	Training Accuracy	Testing Accuracy
DT	0.979	0.445
RF	0.979	0.486
GBC	0.922	0.467
KNN	0.598	0.422
ETC	0.4799	0.472

even though to a slightly lesser degree. The KNN model, with a training accuracy of 0.598 and a testing accuracy of 0.422, reflects its struggle with the higher-dimensionality and noisy data, failing to capture underlying patterns effectively. The ETC shows the lowest training accuracy at 0.4799, yet its testing accuracy of 0.472 suggests a model that is underfitting, unable to learn adequately from the training data to perform well on the test data.

These results highlight a critical challenge with the current dataset, such that the models are either overfitting or underfitting, leading to suboptimal performance. This issue is particularly pronounced given the complexity and diversity of the sensor data. To address this, the data augmentation technique ACGAN is used as a solution to enrich the dataset, providing more varied and representative samples. By artificially expanding the dataset, data augmentation can help mitigate overfitting, improve model generalisation, and ultimately enhance the predictive performance on unseen data.

B. CLASSIFICATION RESULTS ON THE AUGMENTED DATASET

The augmented dataset, comprising 9,000 samples, was divided into an 80-20 train-test split to evaluate the model's performance. The same models and hyperparameters used for the original dataset were applied to this augmented dataset. After successfully training the models on the training set, the testing set was used to evaluate the models' performance. The results of this evaluation are presented in Table 3.

The results of the model evaluation on the augmented dataset presented in Table 3 and visualized in Figure 6, reveal insightful performance metrics across several classification

TABLE 3. Classification results on augmented dataset.

Model	Training Accuracy	Testing Accuracy	Precision	Recall	F1-Score
DT	0.997	0.907	0.91	0.91	0.91
RF	0.997	0.936	0.94	0.94	0.94
GBC	0.978	0.921	0.93	0.92	0.92
KNN	0.89	0.840	0.85	0.84	0.84
ETC	0.997	0.9385	0.94	0.94	0.94

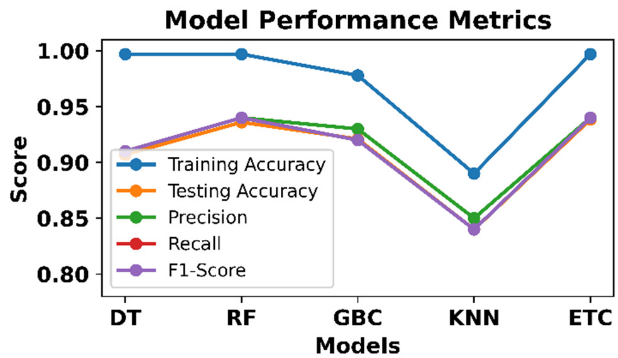


FIGURE 6. Visualization of performance metrices of classifiers on augmented data set.

algorithms. The DT model showed a high training accuracy of 99.7% but a lower testing accuracy of 90.7%, suggesting potential overfitting, with a precision, recall, and F1-score all at 0.91. The RF model performed strongly, with training and testing accuracies of 99.7% and 93.6%, respectively, and precision, recall, and F1-score all at 0.94. Similarly, the GBC had a training accuracy of 97.8%, testing accuracy of 92.1%, precision of 0.93, recall of 0.92, and F1-score of 0.92. The KNN model had lower accuracies of 89% for training and 84% for testing, with a precision of 0.85, recall of 0.84, and F1-score of 0.84. The ETC excelled, achieving 99.7% training accuracy, 93.85% testing accuracy, and precision, recall, and F1-score all at 0.94. The RF and ETC models exhibited the best performance on the testing set, with high testing accuracies and balanced precision, recall, and F1-scores, making them the most effective classifiers for this augmented dataset.

In this manuscript, 10-fold cross-validation was performed to thoroughly assess the reliability and robustness of the models. This method ensures that the performance metrics are not dependent on a single train-test split but are instead averaged over multiple iterations to provide a more comprehensive evaluation. The results, including the average accuracy, and StdDev across all folds, are presented in Table 4 and visualized in Figure 7. This approach helps in validating the consistency and generalisation ability of the models on the augmented dataset.

The 10-fold cross-validation results, presented in Table 4, reveal the performance and reliability of the models. The DT achieved an average accuracy of 91.30% with a standard deviation of 0.01. The RF demonstrated strong performance

TABLE 4. Kfold cross validation results of the classifiers on the augmented dataset.

Model	Average Accuracy ± StdDev
DT	0.9130±0.01
RF	0.9335±0.0057
GBC	0.9260±0.00
KNN	0.8533±0.01
ETC	0.9372±0.00659

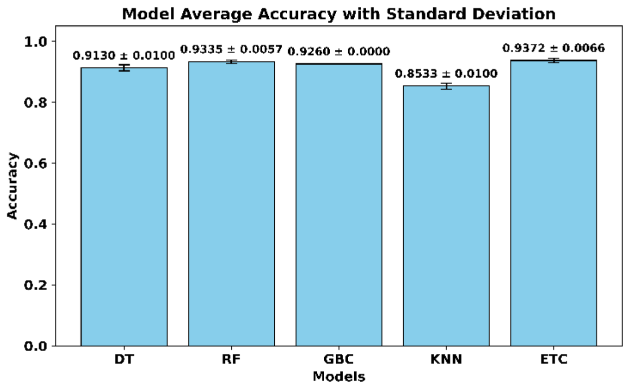


FIGURE 7. Visualization of kfold cross validation results of the classifiers on the augmented dataset.

with an average accuracy of 93.35% and a standard deviation of 0.0057. The GBC showed stable performance with an average accuracy of 92.60% and no reported variability. The KNN model had a lower average accuracy of 85.33% with a standard deviation of 0.01. The ETC outperformed the other models, achieving the highest average accuracy of 93.72% and a standard deviation of 0.00659.

Based on the results presented in Table 3 and Table 4, the ETC is selected as the best model. Its superior average accuracy of 93.72% and relatively low standard deviation of 0.00659 demonstrate its robustness and reliability. The ETC model performed better due to its ability to handle high-dimensional data effectively and its robustness to overfitting, which is achieved through averaging multiple decision trees. This results in high generalisation performance, making it the most reliable model for this dataset.

C. DISCUSSION

The results from this study offer important insights into the performance of different machine learning models for fault detection in power grids, with a particular focus on the impact of data augmentation using ACGAN. Initially, the models trained on the original dataset exhibited significant challenges, primarily due to overfitting and underfitting. The DT and RF models demonstrated high training accuracies (97.9%), but their testing accuracies were much lower, at 0.445 and 0.486, respectively. This large gap indicates that these models memorized the training data, learning noise and specific patterns that did not generalize well to unseen data. The RF model, despite being an ensemble method designed to reduce overfitting, struggled with the variance introduced by the small and imbalanced dataset. On the other hand, the ETC model showed moderate performance, with a training accuracy of 47.99% and testing accuracy of 47.2%, indicating potential underfitting. Although the model was not able to learn complex patterns, its stable performance suggested it was less prone to overfitting than the DT and RF models. The GBC also demonstrated a training accuracy of 92.2%, but its testing accuracy was lower at 46.7%, suggesting some overfitting while still capturing useful patterns.

After applying ACGAN for data augmentation, significant improvements were observed across all models. The augmented dataset, which increased the sample size to 9,000, addressed the data imbalance by generating synthetic data, especially for underrepresented fault types. This not only balanced the dataset but also enhanced the robustness of the models. Following augmentation, the models demonstrated substantial improvements in testing accuracy, with the RF and ETC models achieving testing accuracies of 93.6% and 93.85%, respectively, and balanced performance metrics across precision, recall, and F1-score. The ETC model performed the best with the highest testing accuracy and exhibited low standard deviation during cross-validation, confirming its robustness and generalization ability. The use of ACGAN to generate realistic, class-conditional samples for underrepresented fault types proved crucial in overcoming the challenges posed by the limited and imbalanced dataset. These improvements suggest that ACGAN can significantly mitigate overfitting, enhance the model's ability to generalize, leading to better predictive performance, especially when dealing with complex, imbalanced datasets in fault detection.

The improved results from the augmented dataset also have important implications for real-world power grid applications. The use of spectral and statistical features, which capture both time-domain and frequency-domain signal characteristics, aligns with the types of data commonly collected by sensors in modern power grids. This ensures that the methodology proposed in this study is directly applicable to real-world settings. The ability to handle large datasets and diverse fault types is critical for operational power grids, where fault scenarios are dynamic and can vary widely in both frequency and type. The ACGAN-based augmentation enhances the model's ability to recognize these diverse fault events, making it more suitable for large-scale, real-world deployment. Furthermore, the balanced class representation achieved through synthetic data generation ensures that the model is less likely to be biased towards more common fault types, improving the reliability of the fault detection system in practice.

While the results from the augmented dataset were promising, there are still several factors that influenced model performance. The training of the models with a relatively complex and diverse set of features required careful tuning and parameter optimization. Despite this, ensemble models like RF and ETC performed better than simpler models such as KNN, which struggled with the high-dimensional data, as reflected in its lower performance. The complexity of the sensor data set, with its noise and varying fault types, posed challenges for KNN, which is sensitive to the dimensionality of the feature space. However, the better performance of ensemble models in this case suggests that more sophisticated algorithms are better suited to handle the complexities inherent in power grid fault detection. Additionally, while the use of ACGAN effectively addressed data imbalance, further exploration of more advanced feature extraction methods or deep learning techniques could provide more powerful representations of the fault data, potentially improving detection accuracy even further.

VI. CONCLUSION

This study explores new techniques for strengthening the reliability and stability of modern electricity systems. This study focuses on the advancement of methodologies for identifying, forecasting, and categorizing faults in these crucial networks. Leveraging the GESL dataset, a comprehensive strategy was established in this work. Initially, an extensive array of spectral and statistical features was derived from the data. Furthermore, to tackle the inherent imbalance in the data, an ACGAN for the purpose of data augmentation was used. This strategy efficiently increases the size of the dataset, hence improving the generality of models. After data preparation, a set of machine learning methods, such as Decision Trees, Random Forest, Extra Tree Classifier, Gradient Boosting Classifier, and K-Nearest Neighbors, were trained and evaluated. The Extra Tree Classifier emerged as the most successful model, attaining a testing accuracy of 93.85%. In addition, the strength and reliability of this model were validated using a 10-fold cross-validation technique, resulting in an average accuracy of 93.72% and a remarkably low standard deviation. The results emphasize the significant impact of sophisticated data augmentation approaches, such as ACGAN, on improving the effectiveness of machine learning models for predicting and classifying power grid faults. This research highlights the crucial significance of extensive datasets and creative machine learning methods in strengthening the durability and stability of contemporary power grids, eventually protecting society from the many repercussions of service disruptions.

Future work will aim to incorporate additional event scenarios and explore advanced feature extraction techniques, including wavelet transforms and Hilbert-Huang Transforms, to further enhance fault classification accuracy. The development of real-time fault prediction and classification systems will be a priority to, ensure seamless integration with actual operational power grids. A focus on optimizing models for large-scale deployment in dynamic and complex grid environments will be essential to improve scalability and practical applicability. Moreover, leveraging sensor data fusion and advanced data augmentation strategies can significantly boost detection accuracy, contributing to greater resilience and reliability in modern power grid systems.

REFERENCES

[1] M. Alonso, H. Amaris, D. Alcala, and D. M. Florez R., "Smart sensors for smart grid reliability," *Sensors*, vol. 20, no. 8, p. 2187, Apr. 2020.

[2] Q. N. Minh, V.-H. Nguyen, V. K. Quy, L. A. Ngoc, A. Chehri, and G. Jeon, "Edge computing for IoT-enabled smart grid: The future of energy," *Energies*, vol. 15, no. 17, p. 6140, Aug. 2022.

[3] A. E. L. Rivas and T. Abrão, "Faults in smart grid systems: Monitoring, detection and classification," *Electric Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106602.

[4] M. A. Mahmoud, N. R. M. Nasir, M. Gurunathan, P. Raj, and S. A. Mostafa, "The current state of the art in research on predictive maintenance in smart grid distribution network: Fault's types, causes, and prediction methods—A systematic review," *Energies*, vol. 14, no. 16, p. 5078, Aug. 2021.

[5] A. R. Ekti, A. Wilson, J. Olatt, J. Holliman, S. Yarkan, and P. Fuhr, "A simple and accurate energy-detector-based transient waveform detection for smart grids: Real-world field data performance," *Energies*, vol. 15, no. 22, p. 8367, Nov. 2022.

[6] A. J. Wilson, A. R. Ekti, J. Follum, S. Biswas, C. Annalicia, J.-Y. Joo, O. Aziz, and J. Lian, "The grid event signature library: An open-access repository of power system measurement signatures," *IEEE Access*, vol. 12, pp. 76207–76218, 2024.

[7] M. Morales-Rodriguez, N. Srinivas, J. Olatt, G. Hahn, A. Wilson, P. Fuhr, J. Joo, E. Stewart, and I. Chakraborty, "Big data framework for the development of a signature library for the power grid," Lawrence Livermore Nat. Lab. (LLNL), Livermore, CA, USA, Tech. Rep. LLNL-CONF-817699, 2020.

[8] C. Annalicia and J. Y. Joo, "Grid event signature library analytics report: Signature matching tool development efforts," Lawrence Livermore Nat. Lab. (LLNL), Livermore, CA, USA, Tech. Rep. LLNL-TR-854185, 2023.

[9] O. Alaca, A. R. Ekti, A. Wilson, J. Holliman, E. Piersall, S. Yarkan, and N. Stenvig, "Detection of grid-signal distortions using the spectral correlation function," *IEEE Trans. Smart Grid*, vol. 14, no. 6, pp. 4980–4983, Nov. 2023.

[10] J. D. Follum, S. Biswas, P. V. Etingov, and T. Yin, "A novel method for setting meaningful thresholds for RMS-energy oscillation detectors," Pacific Northwest Nat. Lab. (PNNL), Richland, WA, USA, Tech. Rep. PNNL-SA-185255, 2024.

[11] A. J. Wilson, B. R. J. Warmack, R. A. Kerekes, and P. D. Brukiewa, "Comparison of power system current sensors via playback of electrical disturbances," in *Proc. IEEE/PES Transmiss. Distribution Conf. Expo. (TD)*, Apr. 2022, pp. 1–5.

[12] A. J. Wilson, A. R. Ekti, and Y. Liu, "Power system event detection using the energy detector: A performance analysis," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Jan. 2023, pp. 1–5.

[13] I. Siniosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Fouliras, and P. Sarigiannidis, "A unified deep learning anomaly detection and classification approach for smart grid environments," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 2, pp. 1137–1151, Jun. 2021.

[14] S. Biswas, J. Follum, P. Etingov, X. Fan, and T. Yin, "An open-source library of phasor measurement unit data capturing real bulk power systems behavior," *IEEE Access*, vol. 11, pp. 108852–108863, 2023.

[15] M. N. Fekri, A. M. Ghosh, and K. Grolinger, "Generating energy data for machine learning with recurrent generative adversarial networks," *Energies*, vol. 13, no. 1, p. 130, Dec. 2019.

[16] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Comput. Ind.*, vol. 106, pp. 85–93, Apr. 2019.

[17] R. Wang, S. Zhang, Z. Chen, and W. Li, "Enhanced generative adversarial network for extremely imbalanced fault diagnosis of rotating machine," *Measurement*, vol. 180, Aug. 2021, Art. no. 109467.

[18] M. Farajzadeh-Zanjani, E. Hallaji, R. Razavi-Far, M. Saif, and M. Parvania, "Adversarial semi-supervised learning for diagnosing faults and attacks in power grids," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3468–3478, Jul. 2021.

[19] J. Cho and K. Yoon, "Conditional activation GAN: Improved auxiliary classifier GAN," *IEEE Access*, vol. 8, pp. 216729–216740, 2020.

[20] Q. Sun, F. Peng, X. Yu, and H. Li, "Data augmentation strategy for power inverter fault diagnosis based on Wasserstein distance and auxiliary classification generative adversarial network," *Rel. Eng. Syst. Saf.*, vol. 237, Sep. 2023, Art. no. 109360.

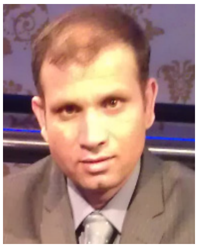
[21] X. Xia, R. Togneri, F. Sohel, and D. Huang, "Auxiliary classifier generative adversarial network with soft labels in imbalanced acoustic event detection," *IEEE Trans. Multimedia*, vol. 21, no. 6, pp. 1359–1371, Jun. 2019.

[22] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018.

[23] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[24] P. K. Saluja and D. Vathana, "Rectifying mode collapse in GANs," in *Proc. Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC)*, 2022, pp. 1718–1722.

[25] K. Zhang, *On Mode Collapse in Generative Adversarial Networks*. Berlin, Germany: Springer, 2021.



HAFEEZ UR REHMAN SIDDIQUI (Student Member, IEEE) received the B.Sc. degree in mathematics from Islamia University Bahawalpur (IUB), and the M.Sc. and Ph.D. degrees in electronic engineering form London South Bank University, in 2012 and 2016, respectively. His research interests include biomedical and energy engineering applications, data recognition, image processing, and system embedded programming IoT-based smart system incorporation with machine learning. He is a Reviewer of IEEE INTERNET OF THINGS JOURNAL.

ROBERT BROWN received the M.Eng. degree in electrical engineering and power systems and the Ph.D. degree in electrical engineering from London South Bank University, U.K., in 2017 and 2024, respectively. He is currently a Research Associate with the School of Engineering, London South Bank University. His main research interests include instrumentation and software for smart electricity grids. He is a member of IET and CIGRE.

ADIL ALI SALEEM received the B.S. degree in computer science from The University of Lahore, Lahore, Pakistan, in 2016, and the M.S. degree in computer science from the Khwaja Fareed University of Engineering and Information and Technology (KFUEIT), Rahim Yar Khan, Pakistan, in 2021, where he is currently pursuing the Ph.D. degree with the Institute of Computer Science. His research interests include the IoT-based smart system embedded with machine learning, text mining, and biomedical engineering.

MUHAMMAD AMJAD RAZA was born in Liāquatpur, Pakistan, in 1992. He received the B.S. degree in information technology from Islamia University Bahawalpur, Pakistan, in 2017, and the M.S. degree in computer science from the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan, in 2021, where he is currently pursuing the Ph.D. degree.

Building on his educational background, he has been actively involved as a Lecturer with Riphah International University Sahiwal, since September 2024, contributing to the academic environment and sharing his knowledge with aspiring IT professionals. Simultaneously, he has been advancing his research endeavors, focusing on the fields of the Internet of Things and data science. Shows his dedication to furthering his expertise and contributing to the academic and research community in his chosen field.

SANDRA DUDLEY (Member, IEEE) was born in Carlow, Ireland. She received the B.Sc. (Hons.) and Ph.D. degrees in physics from the University of Essex, Essex, U.K., in 1998 and 2004, respectively. She spent two years working as a Postdoctoral Researcher with Essex University on a British Telecom Project resulting in world-record broadband-system power consumption for last mile access. She is a Professor of Communication Systems with the School of Engineering, London South Bank University, London, U.K., with research interests in remote signal processing, machine learning approaches in sustainable, smart energy systems and the integration of power and communication infrastructures. She collaborates and leads on several U.K., EU, and IUK research grants with applications towards remote monitoring and data processing.