

## Article

# Intrusion Detection in Smart Power Networks Using Inception-V4 Neural Networks Optimized by Modified Polar Fox Optimization Algorithm for Cyber-Physical Threat Mitigation

Chao Tang \*, Linghao Zhang and Hongli Liu

State Grid Sichuan Electric Power Research Institute, Chengdu 610072, China;  
gwsddl-dkyzlh@sc.sgcc.com.cn (L.Z.); jliu34999@gmail.com (H.L.)

\* Correspondence: lwtg\_tc@163.com

## Abstract

Threats that are caused by cyber-attacks on intelligent power networks promote the implementation of sophisticated intrusion detection devices, which can effectively detect advanced attacks. In this paper, a new model is introduced that combines the Modified Polar Fox Optimization Algorithm (MPFA) with an Inception-V4 deep neural network to enhance the effectiveness of the threat detection task. The MPFA optimizes inception-V4 hyperparameters and architecture to balance the exploration and exploitation processes of the courtship learning process and fitness-based scaling. The optimized model on the smart grid monitoring power is shown to perform well; it achieves over 99.5% accuracy, precision, recall, and F1-score on the detection of various attacks, including False Data Injection, Denial-of-Service, and Load Redistribution, and has a favorable computational overhead, thus it can be considered a formidable solution to protect critical smart grid infrastructure. The optimized model, evaluated on the Smart Grid Monitoring Power dataset, achieves state-of-the-art performance with an accuracy of 99.63%, a precision of 99.61%, a recall of 99.65%, and an F1-score of 99.63% for the detection of various cyber-physical attacks, including False Data Injection, Denial-of-Service, and Load Redistribution. It also maintains a favorable computational overhead, thus presenting a formidable solution for protecting critical smart grid infrastructure.

**Keywords:** intrusion detection; smart grid security; cyber-physical threats; Modified Polar Fox Optimization (MPFA); Inception-V4; hyperparameter optimization



Academic Editor: Ahmed Abu-Siada

Received: 13 November 2025

Revised: 24 December 2025

Accepted: 26 December 2025

Published: 13 January 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

The worldwide energy market is experiencing a paradigm shift under the influence of digitalization and the insertion of renewable energy, which is radically changing the conventional power systems, converting them into intelligent, network-based systems [1]. Modern smart grids now include an extensive number of sensors, advanced communication standards and automated control systems, forming a closely knit ecosystem also known as a Cyber-Physical Power System (CPPS) [2]. It is an architecture of deep integration in which the physical electrical infrastructure, generators, transformers, and transmission lines are constantly observed and controlled by an overlaying digital nervous system [3].

The resulting convergence creates great values, such as increased operational efficiency, predictive maintenance, dynamic demand response, and enhanced grid resiliency [4]. But even the intense mix of information technology (IT) and operational technology (OT)

drastically increases the attack surface, placing the critical national infrastructure at risk of advanced cyber-physical attacks [5]. Where the attacks were mostly hypothetical or had physical access, remote intrusions can now be made by adversaries, which directly interfere with the operations of the grid, potentially causing devastating consequences [6].

Smart grids are exposed to an ever-growing risk of attacks, including False Data Injection (FDI), where attackers steal sensor data to manipulate the control systems. Denial-of-Service (DoS) attacks that destroy critical communications and Load Redirection attacks may cause cascading failures [7]. These intrusions can lead to widespread blackouts, severe equipment damage, and profound economic and societal disruption, underscoring the existential importance of grid cybersecurity [8].

The urgent need for this study stems from the ever-growing number and complexity of cyber-physical attacks on smart grids, which are considered a critical element of national infrastructure [9]. With the increasing digitalization and interconnectivity of power systems, the susceptibility of power systems to remote and coordinated cyberattacks has never been higher, presenting unprecedented threats to grid stability, public safety, and economic security. The development of advanced intrusion detection systems is a current research topic among researchers, as traditional security mechanisms, such as signature-based intrusion detection and rule-based monitoring, are proving insufficient in detecting multi-vector threats that continuously evolve, including False Data Injection and adaptive Denial-of-Service attacks [10]. The intelligent, adaptive detection frameworks are therefore not a random scholarly undertaking, but a dire operational need to guarantee the robustness, dependability, and reliability of next-generation power networks. The proposed study is a direct response to this imperative because it offers a high-fidelity, optimized deep learning model that can identify subtle and complex cyber-physical intrusions in real-time, thereby assisting in the protection of critical energy infrastructure [11].

Intrusion Detection Systems (IDS) are crucial in order to defeat these changing threats. The conventional, signature-based IDS, based on the existing data of known attack patterns, has proven to be inadequate [12]. They do not pick new, so-called zero-day attacks and are not able to adjust to the multi-modal features and subtle reliance of new-age cyber-physical attacks, leaving critical infrastructure in a high-risk situation [13].

In turn, the shift to data-driven methods and, as such, the emergence of the Machine Learning (ML) and Deep Learning (DL) paradigm [14] have become the focus of research. Their capability to learn discriminative patterns of complexity directly out of enormous data streams of operational information provides opportunities for adaptive intelligent threat detection that can adapt to the dynamic threat environment [15]. Such methods may reveal the hidden anomalies that could not be detected in rule-based systems [16].

In spite of this pledge, there are significant challenges to the use of traditional ML and DL models in the field of smart grid security [5]. Models, including Support Vector Machines (SVMs) or basic Convolutional Neural Networks (CNNs), often exhibit unstable behavior and poor extrapolation to various attack cases due to the high-dimensional and spatiotemporal characteristics of cyber-physical data [17]. Moreover, they are susceptible to the labor-intensive manual process of hyperparameter tuning and feature engineering, which cannot be both practical and optimal in the dynamic grid setting [18].

This highlights one of the most significant research gaps: the need for an automatically optimized, high-capacity detection structure. A perfect solution should have the ability to adapt to the specifics of power system data automatically, extracting complex features from raw and combined cyber-physical data streams, and provide specific and robust threat classification without human assistance.

To address this gap, this paper presents an original hybrid intrusion detection framework that strategically integrates the deep hierarchical feature extraction strength of the

Inception-V4 neural network with a customized Modified Polar Fox Optimization Algorithm (MPFA). We do not play an architectural design role, but rather in the development and use of MPFA to methodically and automatically optimize hyperparameters and the structure of the Inception-V4 model. This is a highly optimized system, which is particularly fidelity-intensive and sensitive to stationary cyber-physical threats, such as FDI, DoS attacks, and Load Redistribution attacks, with the intention of delivering a robust, intelligent layer of protection to next-generation power infrastructures.

## 2. Literature Review

The increased digitization of modern power systems has transformed the traditional electrical grid into highly interdependent cyber-physical smart grids. Though the transformation will facilitate the achievement of such potent functionalities as real-time monitoring, demand response, integration of the distributed energy resources, and remotely controlled resources, it will also put power infrastructures at risk of a broader range of cybersecurity threats. It can be applied maliciously, such as in data injection, denial-of-service attacks, and tampering attacks, which can compromise measurement data, disrupt the integrity of intelligent electronic devices (IEDs), and compromise the stability of the grid. Due to the inability of the classical protection approaches to recognize this sophisticated form of intrusion, machine learning (ML) and deep learning methods can be successfully utilized as tools to detect suspicious objects and thwart cybercrime in smart grids. These approaches will enhance situational awareness, facilitate quick response mechanisms, and provide reliability, resilience, and security for power systems operating in increasingly complex cyber-physical environments.

Sadi et al. [19] research on cybersecurity of smart inverters utilized in distributed energy resources, which is vital to cloud computing, remote monitoring and peer-to-peer energy trading in the present power systems. After appreciating the severe threat of data injection attacks that have the potential to alter the properties of measurements and destabilize the grid, they provided a time-driven machine learning-based anomaly detection framework to identify cyber intrusions to sets of control signals and DC voltage measurement biasing in voltage source convertors (VSCs) in wind generators. The paper discussed the impacts of four categories of significant attacks on smart VSCs and wind farms, such as denial-of-service, tampering, stealthy, and data intrusion. Another set of time-sequence machine-learned intrusion detection was developed and compared to the autoencoders and clustering-based models of intrusion detection. The framework's performance has been tested using the IEEE 39-bus power system, which features four wind farms positioned in different areas. The results demonstrated the efficiency and soundness of the proposed model in detecting cyberattacks in innovative VSC systems, utilizing multiple performance measures.

Sahani et al. [20] conducted an extensive survey of machine learning (ML)-based intrusion detection systems (IDSs) in an innovative grid environment, with a specific focus on their greater applicability in ensuring enhanced system security against new cyber threats. Although the application of ML-based IDS techniques in general computing systems has greatly enhanced network defense, their application in smart grids is relatively underutilized, making the grid more prone to attack due to the prevalence of common network structures. The study article investigated the use of ML-based IDS in transit and distribution of smart grids, considering the potential of dealing with context-related security risks. Furthermore, it also mentioned the development of datasets and how they were used to train the IDS model, compared various ML algorithms used in the literature survey, and analyzed key performance measures, including training performance and testbed results. The authors also provided insights, challenges, and a summary of future

directions on how a more robust, adaptive, and interpretable ML-based IDS structure can be created to strengthen innovative grid cybersecurity.

Aljohani et al. [21] introduced an intrusion detection and mitigation system (IDMS), which operates on the basis of deep learning neural networks (DLNNs) and will be employed to enhance the safety of a digitalized power system. As more and more cyber and physical infrastructure is incorporated into the smart grid, the threats of cyberattacks increase, and cybercriminals can inject fake data that can cause unneeded protective actions and cause a mass outage. To overcome this problem, the current paper proposes a framework based on the DLNN, which can detect, classify, and identify intrusions in the smart grid. The system identifies the disturbances first and isolates a one-point and coordinated attack. It thereby isolates the compromised intelligent electronic device (IED) and forecasts its current waveform using a long short-term memory model (LSTM) to ensure the system's observability in the future. The created IDMS was introduced on a modified IEEE 13-bus system, and the simulation test results showed high precision in intrusion detection, classification, localization, and forecasting, which proves the possibility of the high efficiency of the developed IDMS in protecting operations in intelligent grids.

Ankitdeshpandey et al. [22] investigated the use of machine learning (ML) algorithms in cyberattack detection and identification in smart power grids that are currently vulnerable to cyberattacks due to their connection to the Internet. The article utilized data from MSU-ONL at Mississippi State University and Oak Ridge National Laboratories to construct a deep neural network (DNN) model that classifies data into three categories: power system attack, normal, and no-event. OneR, K-Nearest Neighbor (KNN), Random Forest, Support Vector Machine (SVM), and Naive Bayes are also a set of conventional ML methods that were used and compared to evaluate the performance of DNN in the detection of intrusions. Principal Component Analysis (PCA) was used to reduce the data dimensions in order to establish its influence on the model performance. The empirical results indicated that the Random Forest model was the most precise in the attack detection, SVM and DNN scored higher than the PCA model. It was also determined in the results that the SVM, Random Forest, and DNN algorithms can be used to deploy the intrusion detection systems (IDS) to power grid cybersecurity.

Li et al. [23] addressed the increasing cybersecurity concerns of modern smart grids, the implementation of which is premised on sound cyber-physical connectivity to condition supervision, and is prone to various cyberattacks. They suggested an Adaptive Deep Learning (ADL) framework that consists of three modules: data pre-processing, neural network pre-training, and classification, to optimize the performance of the available machine learning-based intrusion detection classifiers. The ADL algorithm was used to calculate the optimum number of layers and the number of neurons per layer, depending on the characteristic dimension of the network traffic data. Transfer learning enabled it to obtain new abstract features in the original high-dimensional data. In such a way, it is, literally, a combination of a deep understanding and a conventional machine learning method. The NSL-KDD data was used to train the algorithm, and the results of the experiments demonstrated that the proposed ADL model achieved relatively higher classification rates and required less training compared to existing models, which highlighted the developmental potential of network security in smart grids.

Cavus et al. [24] proposed a cyber-resilient data-driven optimization system of real-time energy operation of EV-integrated smart grids. The framework, which integrates genetic algorithms and reinforcement learning with real-time analytics, can schedule EV charging in an adaptive manner based on dynamic electricity pricing, mobility patterns, and grid load variability. It is the first to combine adaptive optimization, resilient forecasting in incomplete data (MAE of 0.25 kWh, MAPE of less than 20% even with 25% of data missing),

and a lightweight blockchain-inspired security protocol with an intrusion detection system (94.1% accuracy, AUC of 0.97 and fast attack detection). On European data of a smart grid, the strategy minimally decreased daily peak demand (9.6 percent), more evenly distributed charging load (peak normalised utilisation fell to 0.7 and kept 0.4 s) and continued to optimize at run times of less than 0.4 s on a large scale. The best forecasting model (RMSE: 0.853 kWh) was CatBoost. Another extension of the location-based charging infrastructure (LOSC) planning to a conceptual one proposed by the research is to plan deployment in line with predicted demand. In general, the framework has a high level of technical strength, operationality, and scalability for the intelligent EV-grid systems in the future. Table 1 shows a comparative analysis of Intrusion Detection Techniques to Smart Grid security.

**Table 1.** Comparative Analysis of Intrusion Detection Techniques for Smart Grid Security.

| Method/Study                                                 | Core Technique                                                                  | Key Advantages                                                                                               | Reported Performance                                                                                                                        | Main Limitations                                                                                                                                                 |
|--------------------------------------------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Signature-based & Conventional ML (e.g., SVM, Random Forest) | Rule matching, statistical classifiers.                                         | Interpretable, low computational cost for known threats.                                                     | Varies; ~90–96% on balanced datasets.                                                                                                       | Fail against novel “zero-day” attacks; require manual feature engineering; poor adaptation to evolving threats.                                                  |
| Autoencoder-based Anomaly Detection                          | Unsupervised reconstruction error.                                              | Can detect novel attacks without labeled malicious data.                                                     | High recall for significant anomalies.                                                                                                      | Struggles with subtle, multi-modal attack signatures (e.g., stealthy FDI); high false positives on noisy grid data.                                              |
| Graph Neural Networks (GNNs)                                 | Models grid topology as a graph.                                                | Excels at detecting structural attacks (e.g., Load Redistribution).                                          | High accuracy when topology is known.                                                                                                       | Requires precise, complete, and often static topological knowledge; less effective on pure cyber-attacks (DoS).                                                  |
| Hybrid GA/RL with IDS ([5])                                  | Genetic Algorithm + Reinforcement Learning for scheduling, with a separate IDS. | Integrates optimization with security; adaptive scheduling.                                                  | IDS Accuracy: 94.1%, AUC: 0.97, Fast detection (~50–300 ms).                                                                                | Accuracy notably lower than DL-focused IDS; separate modules for scheduling and intrusion detection.                                                             |
| DLNN/LSTM-based IDMS ([21])                                  | Deep Learning Neural Networks, Long Short-Term Memory.                          | Suitable for time-series data; can classify and locate intrusions.                                           | High precision in reported studies.                                                                                                         | May not capture multi-scale spatial features; performance reliant on architecture tuning.                                                                        |
| EV energy management framework ([24])                        | (GA+ RL) + Lightweight blockchain-inspired security IDS + CatBoost              | Efficient, Adaptive scheduling under uncertainty                                                             | -9.6% daily peak demand reduction<br>-27% less energy delivered at original peak hour<br>-~25% increase in demand<br>-Forecasting MAE: 0.25 | -not implemented—<br>Framework validated only on European datasets<br>-Cybersecurity module is not full blockchain; may lack formal decentralization guarantees) |
| Proposed MPFA-Optimized Inception-V4 (This work)             | Metaheuristic-optimized deep hierarchical CNN.                                  | Automatic hyperparameter tuning, superior multi-scale feature extraction, balanced exploration/exploitation. | Accuracy: 99.63%, Precision: 99.61%, Recall: 99.65%, F1-Score: 99.63%                                                                       | Offline optimization phase adds initial computational overhead.                                                                                                  |

Other notable paradigms for smart grid IDS systems, besides the supervised and deep learning models covered, include autoencoder-based anomaly detection systems and graph-based systems. Autoencoders, which are trained to recover normal functioning data, also do a good job of identifying anomalies (e.g., FDI attacks) as those with a significant reconstruction error, and thus can be beneficial in identifying novel, previously unseen

attacks without the need for labeled malicious data. On the same note, graph-based IDS represent the physical topology and communication network of the smart grid as a graph, and identify intrusions by structural dependencies between the innovative grid components, using Graph Neural Networks (GNNs) to learn the structural dependencies among these components (e.g., load redistribution). Although these are effective, autoencoders can be sensitive to the high dimensionality and multi-modality of cyber-physical data, where attack patterns can be low-level. Graph-based analysis requires accuracy and complete topological information, which is not always available and can also change. The Inception-V4 framework proposed as the optimization of a more efficient MPFA tends to fill in these gaps by utilizing a deep hierarchical network to learn the complex high-level representation of raw integrated cyber-physical data, without the explicit graph modeling, and through the application of metaheuristic optimization to adapt the model to high-fidelity detection in a variety of known attack categories, assuring robustness and high accuracy in the environment with a clear understanding of the threat landscapes.

The accelerating integration of cyberspace and physical systems in modern innovative power systems has led to the emergence of Cyber-Physical Power Systems (CPS), which has resulted in a significant increase in grid efficiency, situational awareness, and operational flexibility. Such integration, however, also creates significant vulnerabilities to high-tech cyber-physical attacks, including false data injection, denial-of-service attacks, and load redistribution, which can destabilize the grid's operation and cause cascading failures. Traditional intrusion detection systems (IDS), such as rule-based and shallow machine learning approaches, have been found to be ineffective at identifying zero-day attacks and capturing the complex spatiotemporal relationships that smart grid telemetry data imply. This weakness underscores the urgent need for adaptive and intelligent detection infrastructures that can detect threats in real-time with fidelity in dynamic and high-dimensional environments.

To overcome the given challenges, the present study proposes an innovative hybrid framework that combines the Inception-V4 deep neural network with a Modified Polar Fox Optimization Algorithm (MPFOA). It is essential to explain that, even though the Inception-V4 architecture is maintained, meaning that its structure and design, including the multi-branch hierarchical convoluted structure as described in previous literature, are kept, its usage is optimized creatively to detect intrusions in smart grids. The fundamental novelty of this work lies neither in architectural redesign nor in the development and implementation of MPFOA to systematically optimize both the hyperparameters and architectural settings of Inception-V4 for the particular domain of cyber-physical threat detection. To strengthen the performance of the model, MPFOA applies adaptive search, convergence-aware randomization, and courtship-based learning, allowing a tuned Inception-V4 network to informatively extract high-quality spatio-temporal features from smart grid data.

The chief novelties of this work and contributions are as follows:

- **New Optimization Algorithm:** The Modified Polar Fox Optimization Algorithm (MPFA) that utilizes gender-based learning in courtship, adaptive attraction, and fitness-based scaling to enhance convergence and prevent local optima in a large-dimensional search space was developed.
- **First Integration with Inception-V4:** The initial execution of MPFA to fully optimize the hyperparameters as well as the architecture of the Inception-V4 deep neural network directly on cyber-physical intrusion detection in smart grids. Further development of this framework is crucial to enhance its ability to monitor and improve a broad range of variables.
- **Holistic Detection Framework:** A unified optimization framework that simultaneously balances multiple objectives: detection accuracy (precision, recall), model complexity

(parameter count), and computational efficiency, ensuring practical deployability in real-time grid environments.

- **Superior Feature Extraction:** Extraction of a multi-branch hierarchical structure of Inception-V4 to autonomously extract complex spatiotemporal features of integrated cyber-physical data streams in a feature-engineering process-free manner. Results of this work can thus be summarized as follows: (1) The design of the MPFOA to optimize deep neural networks, (2) the design of a specific Intrusion detector system using smart grids, (3) the verification of the framework proposed by the study in relation to existing methods, (4) the study of the interaction between optimization-feature-class mismatch and (5) practical suggestions on how to realize this technology. All these developments together contribute to the evolution of innovative and resilient cybersecurity of power infrastructure in the future generation. The remainder of this paper will be organized as follows. Section 2 presents a comprehensive literature review of smart grid cybersecurity, intrusion detection systems, and optimization methods related to the latter. Section 3 describes the methodology, including the dataset description, the proposed Modified Polar Fox Optimization Algorithm (MPFA), the Inception-V4 architecture, and the incorporated optimization framework. Section 4 presents the experimental setup, results, and comparative analyses, including convergence behavior, classification performances in various types of attacks, ablation studies, and computational trade-offs. Lastly, Section 5 presents the conclusion of the paper, where the main findings are summarized, the implications of the work are discussed, and recommendations are given for future research.

### 3. Method and Materials

Figure 1 shows the combination of electrical grid and cyber systems in a Cyber-Physical Power System (CPPS) with the emphasized awareness of how physical power infrastructure (generators, transformers, transmission lines, and intelligent electronic devices (IEDs)) is seamlessly interconnected with the digital cyber layer through communication networks and supervisory control systems such as SCADA. This intersection enhances grid efficiency, reliability, and situational awareness, but also creates new points of vulnerability to cyber-physical attacks, underscoring the need for further development of intrusion detection mechanisms to ensure the integrity and stability of current innovative power systems.

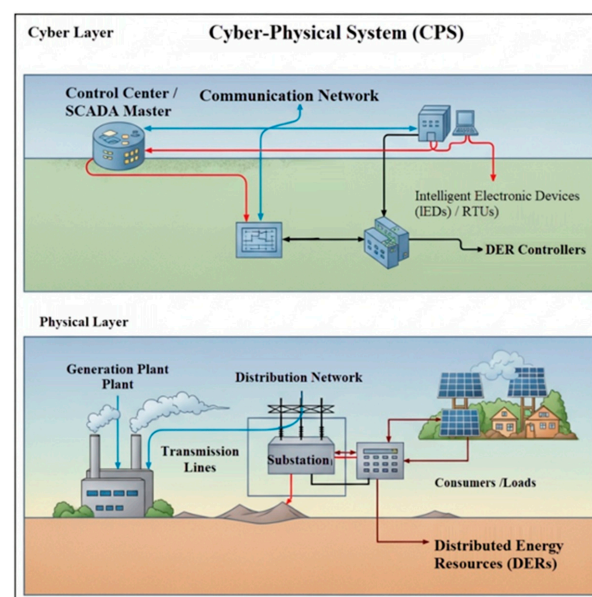


**Figure 1.** A Cyber-Physical Power System (CPPS) that incorporates the electrical grid and cyber systems.

### 3.1. Dataset Description

Figure 2 data reveal that reported cases of cyber-attacks on power systems have been increasing since 2015 through 2025, depending on the dynamics in reported instances of publicly available industrial control system (ICS) incident reporting sources, such as ICS-CERT. The graph above illustrates an upward trend, indicating that the number of cyber intrusion cases has increased by a factor of three over the past decade. This figure graphically illustrates that modern power infrastructures have become increasingly vulnerable, and therefore, intrusion detection systems should be enhanced, such as the proposed Inception-V4 network with the Modified Polar Fox Optimization Algorithm (MPFOA) applied to the network, in an effort to counter new cyber-physical attacks on the network.

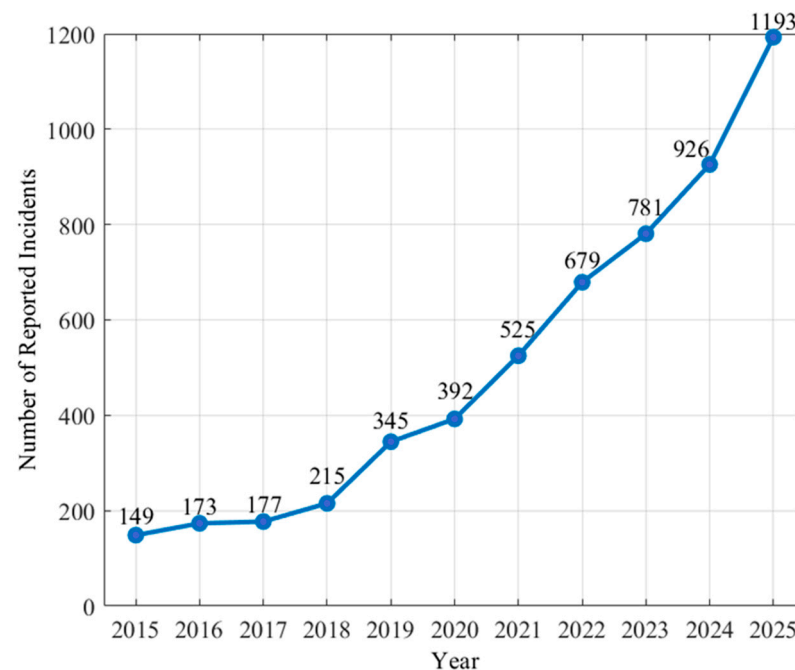


**Figure 2.** Trend of Cyber-Attack Incidents in Power Systems (2015–2025).

The data utilized in the study is the smart grid monitoring power dataset, which is freely available on Kaggle (<https://www.kaggle.com/datasets/bachirbarika/power-system>, accessed on 25 December 2025). This dataset is a multifaceted set of data that describes the behavioral patterns of an energy system incorporating intelligent technologies. This dataset has been chosen due to its comprehensive coverage of contemporary smart grid operations. It offers a flexible range of data that records simultaneous electrical values (e.g., voltage, current, active/reactive power, frequency) and communication network values (e.g., packet delay, packet loss), unlike synthetic or narrowly-scoped collections. This combined space of cyber-physical features is vital for creating an IDS that can identify attacks occurring in both realms. Beyond that, the dataset contains categorized examples of the types of critical attacks applicable to CPPS, including False Data Injection (FDI), Denial-of-Service (DoS), and Load Redistribution, providing a realistic point of reference for testing detection performance against known attack threats. This specific dataset was pre-screened to represent a range of normal operational parameters, as well as different types of cyber-physical intrusions; hence, this particular dataset is extremely useful in intrusion detection research for modern power systems.

The population sample will consist of 37,500 individuals, and the dataset will include 16 predictor variables, comprising both numerical and non-numerical electrical and network parameters. These are line voltage (V), current (I), and active power (P), reactive

power (Q), frequency (f), demand of load, and a binary flag of attack (0), whether it is the normal condition (0) or an intrusion incident (1). The raw data analysis involved an initial examination of the data, which included identifying missing and abnormal values. Unspecified records and records whose value was zero were deleted to preserve integrity. The next step was Min-Max normalization, which was used as a feature scaling, and it looks as follows:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where  $x'$  is a normalized feature,  $x_{min}$ ,  $x_{max}$  are the minimum and maximum values respectively. Normalization works with values that are either 0 or 1, which enables neural networks to be trained faster. An analysis of the classes' distribution showed that neither a sample is balanced, with 78 percent being the standard sample and 22 percent being the attack sample. To address this issue and reduce the bias of the majority in the model, the Synthetic Minority Over-sampling Technique (SMOTE) was employed to generate artificial samples of the minority (attack) group, thereby achieving an almost equal ratio. Moreover, a Pearson correlation analysis was conducted to examine the relationship between features and attribute redundancy, as well as weakly informative features. The correlation coefficients, whose value is less than 0.1 for a feature, were excluded from the training to minimize model efficiency and overfitting. The numbers in Figure 3 show the amount of data in both standard and attack sets in the pre-SMOTE case and in the post-SMOTE case, where the synthetic augmentation has improved the value of balance. Figure 4 graphically presents the scale-location of the chosen features, and the current, power, and voltage variables are highly dependent, which is physically possible as explained by Ohm's Law of electricity and the law of power.

Figure 3 shows the distribution of the classes of the dataset before and after the application of the Synthetic Minority Oversampling Technique (SMOTE). The left panel represents the original, strongly skewed data, in which the Normal type (29,250 samples) displays a significant variance when compared to the Attack type (8250 samples), suggesting that 78 percent of the samples were normal and 22 percent were attack. The right-hand plot presents the balanced dataset, which has been obtained after SMOTE oversampling, where the number of samples in each class is 29,250. This graph is a vivid representation of the impact of SMOTE in reducing the problem of class imbalance by forming synthetic minority samples, thereby preventing the proposed Inception-V4-MPFOA intrusion detector model's data fragmentation process from favoring the auditory capabilities, making it detect its prey depictional discriminative descriptions of normal and malicious occurrences, and hence increasing its generalization and detection capabilities.

Figure 4 indicates the heatmap of the correlation of the features of the smart grid monitoring power dataset. The heat map can be used to visualize the potential pairwise Pearson correlation coefficients of the electrical, operational, and communication variables. The intensity of the heat map ranges from  $-1$  (blue, indicating a strong negative relationship) to  $+1$  (dark red, indicating a strong positive relationship). As it was revealed, Voltage, Current, Active Power (P), Reactive Power (Q), and Temperature are highly inter-correlated with the coefficients being close to 1.00, which means that the above parameters are interrelated in the dynamics of power flow: the larger the current and voltage, the larger the power and temperature in the network would be. Quite to the contrary, the Frequency, Load Demand, Delay, and Loss characteristics are poorly or nearly uncorrelated with electrical quantities, implying that they are independent variables in grid operation and cyber communication behavior. The correlation analysis also confirms that the dataset is composed of highly correlated physical parameters and weakly correlated cyber-related indicators, which provides an abundance of cross-domain relationship multimodal feature space

that the proposed Inception-V4-MPFOA intrusion detection model can exploit to learn cross-domain relationships. The summarized dataset statistics are indicated in Table 2.

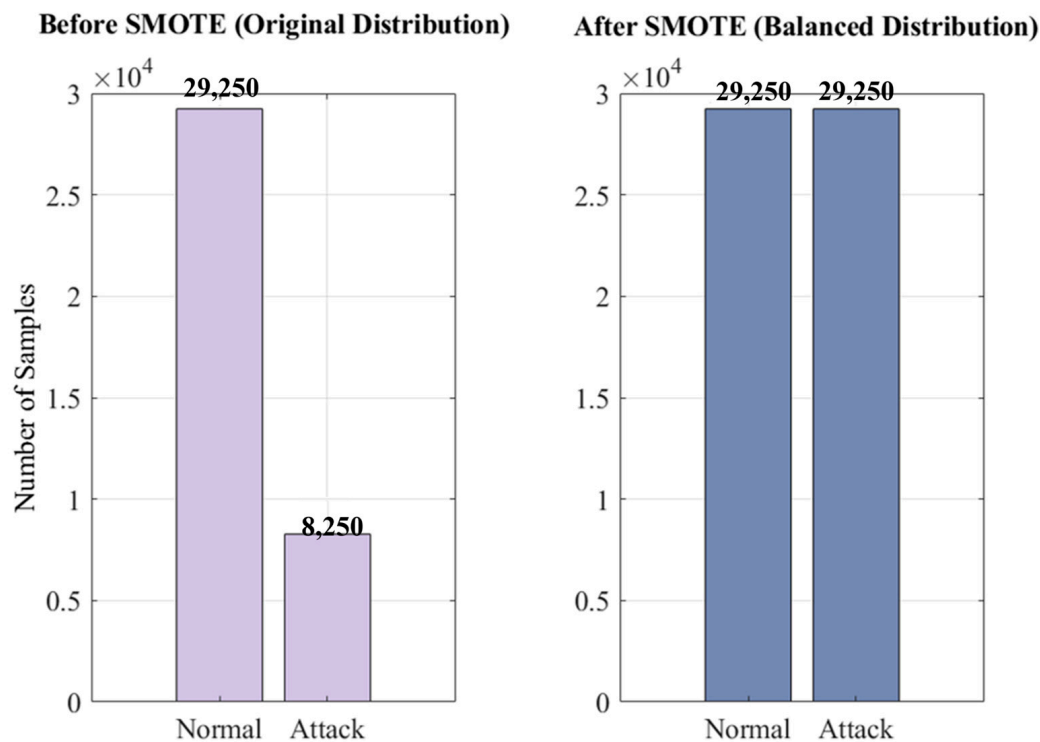


Figure 3. Distribution of Samples Before and After SMOTE.

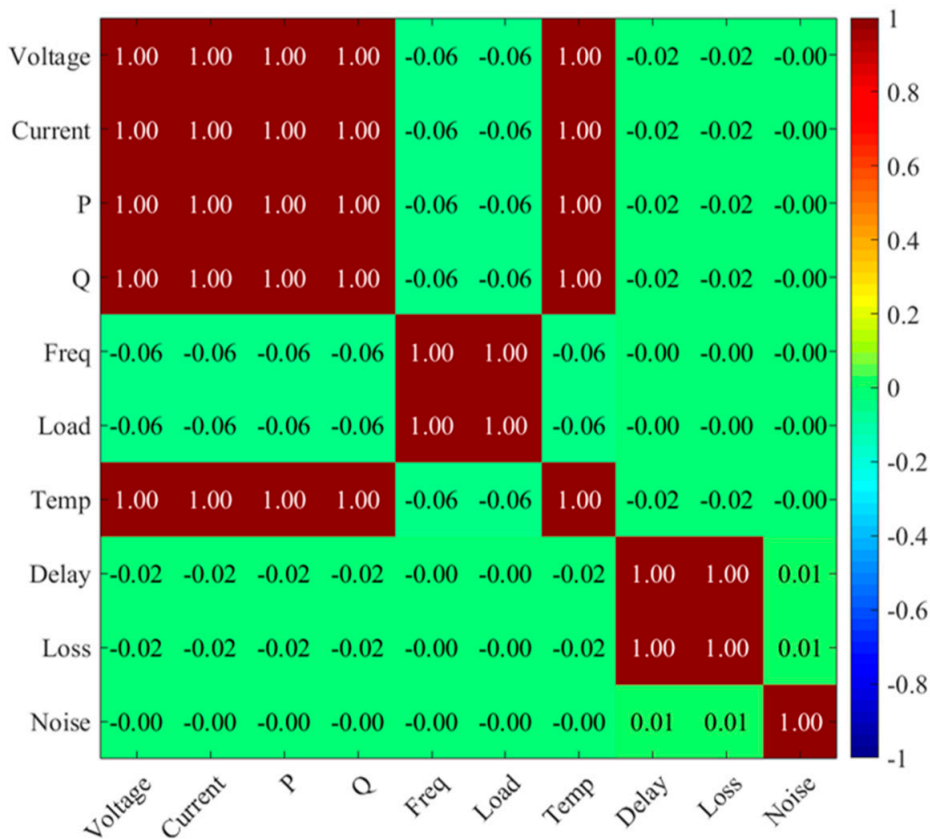


Figure 4. Feature Correlation Heatmap.

**Table 2.** Dataset summary statistics.

| Property           | Value                  |
|--------------------|------------------------|
| Total samples      | 37,500                 |
| Number of features | 16                     |
| Attack categories  | 2 (Normal, Attack)     |
| Missing values     | 0.02%                  |
| Data type          | Mixed (float, integer) |
| Sampling interval  | 1 s                    |
| File format        | CSV                    |

Additionally, Table 3 presents feature descriptions and measurement units.

**Table 3.** Feature descriptions and measurement units.

| Feature Name              | Description                                                            | Unit        |
|---------------------------|------------------------------------------------------------------------|-------------|
| Voltage_V                 | RMS line voltage                                                       | Volts (V)   |
| Current_A                 | Line current                                                           | Amperes (A) |
| Active_Power_P            | Real power flow                                                        | kW          |
| Reactive_Power_Q          | Reactive power flow                                                    | kVAR        |
| Frequency_f               | System frequency                                                       | Hz          |
| Load_Demand               | Power consumption of the load                                          | kW          |
| Bus_Voltage_Deviation     | Voltage deviation from the nominal value                               | %           |
| Line_Temperature          | Thermal state of the line                                              | °C          |
| Power_Factor              | Cosine of phase angle                                                  | –           |
| Total_Harmonic_Distortion | Harmonic content in a signal                                           | %           |
| Packet_Delay              | Network latency                                                        | ms          |
| Packet_Loss               | Data loss percentage                                                   | %           |
| Node_ID                   | a superior auditory sense, thus enabling it to detect its prey despite | Integer     |
| Time_Stamp                | Sampling timestamp                                                     | s           |
| Attack_Type               | Encoded attack category (0/1)                                          | –           |
| Attack_Flag               | Binary label (normal = 0, attack = 1)                                  | –           |

Based on this comprehensive preprocessing and analysis pipeline, smart grid monitoring power will provide a solid foundation that serves as the training data for the proposed Inception-V4-MPFOA intrusion detection model. The space of balanced and normalized features would facilitate significant convergence and accurate learning of the compound interdependency of cyber-physical events, which is characteristic of the current power system's nature.

### 3.2. Modified Polar Fox Optimization Algorithm (MPFOA)

The Modified Polar Fox Optimization Algorithm (MPFOA) is a metaheuristic population-based algorithm that seeks to strike a balance between the complexity of the search space and the exploration-exploitation trade-off. It is based on the previous Polar Fox Optimization (PFO) algorithm, which is enhanced by incorporating a gender-conscious phase of courtship learning and adaptive attraction representation. This modification enhances its efficiency in preventing local optima and achieving efficient convergence behavior.

### 3.2.1. Polar Fox Leash Generation

Such candidates are referred to as a leash or skulk, a social grouping of litter, a group of persons aiding it, and a mating pair [25]. During spring, the group is focused on securing a place to stay and the resulting procreation of their potential young people [26]. To model the set of candidates, the optimizer will begin with a population of these individuals in the solution space distributed randomly and then the following formulas [27]:

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_d^1 \\ x_1^2 & x_2^2 & \dots & x_d^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^k & x_2^k & \dots & x_d^k \end{bmatrix} \quad (2)$$

$$x_j^i = LB + \vec{r}_1 \times (UB - LB) \quad (3)$$

$$LB = [lb_1 \quad lb_2 \quad \dots \quad lb_d] \quad (4)$$

$$UB = [ub_1 \quad ub_2 \quad \dots \quad ub_d] \quad (5)$$

where  $X$  is employed, a candidate with a lower cost value number  $i^{th}$  and a value of  $j$  has been denoted by  $x_j^i$ , number of candidates has been denoted by  $k$ , number of dimensions as depicted by  $d$  and the stochastic vector has been denoted by  $\vec{r}_1$  which is the range of 0–1. In addition,  $LB$  and  $UB$  In turn, they demonstrate the lower and upper limits.

### 3.2.2. Grouping Polar Fox

Each of the groups has a different low fitness value; therefore, they exhibit a larger follow-the-leader effect. Some candidates utilize their knowledge, while others prefer to discover things on their own, and some individuals are very industrious. As a result, the team is categorized into four groups, with members being energized by the leader. They may become somewhat fatigued following this process, which has been significantly shortened into G1i, G2i, G3i, and G4i, respectively. To start with, there is an equal distribution of candidates among the groups. The group on the other hand, retains candidates. In case the target is conveniently hunted, the weights of the groups have been adjusted as follows:

$$W_i^{new} = W_i + \frac{t^2}{NG_i} \quad (6)$$

where, mass of the group  $i$  has been represented as  $W_i$ , the group size of the group  $i$  as  $NG_i$  and the iteration on which one is working as  $t$ . This can be minimized by having an initial value for the weight of the groups.

#### - Experience-based stage

The inhabitants do not passively occupy the months of winter, but lead a gossiping and nomadic life; they attend in small parties in search of food. Equations (6) and (7). In order to imitate the manner in which the candidates hunt, (7) are advised to jump.  $D$  and  $P$  These denote the force and direction of jumping, respectively. The variables can move the individuals who were at their previous position  $x^i(t)$  to the new position  $x^i(t+1)$ . The current process has been estimated by utilizing the following formula:

$$\begin{aligned} x^i(t+1) &= x^i(t) + P \times D \\ PF_i &= PF \times a^{z-1} \\ P &= \vec{r}_2 \times PF_i \\ D &= \cos(\vec{r}_3) \end{aligned} \quad (7)$$

where; the strength aspect of the research of the candidate  $i$  has been modelled by  $PF_i$ , the present number of iterations is denoted by  $t$ , the stochastic vectors of the individuals have been represented by  $\vec{r}_3$  and  $\vec{r}_2$  which are respectively in  $[0, 180]$  and  $[0, 1]$ . The repetitions are given using  $z$ . The process is stopped when a fantastic value of the goal has been reached, and the amount of energy of people is brought to the previously set percentage.

$$\begin{cases} PF_i < m \times PF \\ f(t) < f(t-1) \end{cases} \quad (8)$$

where, the objective function is denoted by  $f$ .

#### - Leader-based stage

A leader has been assigned to each group to achieve the group's objectives. In this case, the position of the leader has been illustrated through  $L$  and has been regarded as a significant objective value. Also, the people shift the position and begin to move  $x^i(t)$  to  $x^i(t+1)$ . Due to the move, the leader's situation has changed. The existing process has been demonstrated in the following:

$$\begin{aligned} x^i(t+1) &= x^i(t) + \vec{r}_4 \times (x^i(t) - L) \times LF_i \\ LF_i &= LF \times b^{y-1} \end{aligned} \quad (9)$$

In which  $LF_i$  is the strength factor,  $y$  is the number of repeats of this stage, and  $\vec{r}_4$  is the stochastic factor, taking values of  $-1$  and  $1$ . When an excellent result has been obtained, and the energy rate of the individuals is lower than the percentage calculated beforehand, then the process stops. The following way has been used to represent the situation:

$$\begin{cases} LF_i < m \times LF \\ f(t) < f(t-1) \end{cases} \quad (10)$$

#### - Leader motivation stage

The candidates cannot initially locate the object through skills. The leader then motivates the candidates. Then, the applicants move randomly between two locations. Therefore, several behavior matrices are obtained, e.g.,  $G4m$ ,  $G3m$ ,  $G2m$ , and  $G1m$ . Thereafter, it leads to more endeavors in small doses dubbed as MLR. In the end, all the individuals are moved to the opposite, which is decided as follows:

$$critical = (NLM > MLM) \text{ or } (t > 0.8 \times NI) \quad (11)$$

where, *critical* As a part to specify that an optimizer has fallen into a local optimum or is near the ultimate point of implementation.

#### - Mutation step

It is known that a tremendous number of such animals die. The young people are deserted mainly by their parents, and the young people murder their brothers or their brothers are brutally. One out of five destroyers of these people is the famous killer disease referred to as rabies. This is discovered to occur at a specific period when the immunity of the candidates has been compromised as a result of nutritional deficiency. At this stage, the less profitable have been replaced with a few new applications to create a stronger group by utilizing the equations as illustrated below.

$$x^i = LB + \vec{r}_5 \times (UB - LB), \quad i = [1, 2, \dots, NM] \quad (12)$$

$$NM = \begin{cases} NP - 1 & \text{if critical} \\ MF \times NP & \end{cases} \quad (13)$$

where  $\vec{r}_5$  is a stochastic vector as depicted in 0–1.

- Fatigue simulation

When the people are motivated by their leader, they will be somewhat exhausted by the time G1r, G2r, G3r, and G4r complete the iterations. Eventually, the behavior matrix is narrowed down to G1i, G2i, G3i, and G4i. Moreover, the fewer the number of individuals in a group that makes less than 10 percent of the entire population, the higher their energy levels increase. Such cases can be explained in terms of Equation (14).

$$Gk = \min(Gk - Gkr, Gki) \quad (14)$$

where,  $k = 1, 2, \dots, 4$ .

### 3.2.3. Modified Polar Fox Optimizer (MPFO)

The original Polar Fox Optimization (PFO) algorithm lacks a mechanism for differentiation between genders within a population, which may be detrimental to its functionality. The algorithm relies heavily on relations and data exchange, but it can be advanced by adding data related to gender. To fill this gap, the present study develops a courtship learning model that utilizes the capacity of the key stone polar fox to learn from the female polar fox, and consequently, to have a more effective search process on earth. A randomization probability is used for the female polar fox in this process, allowing the keystone polar fox to select a female polar fox from the archive, which makes the algorithm more efficient, as suggested by the Courtship Learning (CL) method. To help maximize the performance of the algorithm, the proposed CL methodology has four major characteristics that are recommended for utilization.

- Scaling mechanism

A better alternative is a polar fox with a lower cost value. When properly arranged information about one of the female polar foxes in the archive is employed, a candidate with a lower cost value is more likely to be selected. To achieve this concept, a scaling mechanism has been used for all females. The fact of the transformation of the polar fox can be stated in the following way:

$$M_i = \frac{1}{f(x^i)} \quad (15)$$

$f(x^i)$  in this equation represents the fitness of the  $i^{th}$  female candidate. A female candidate with a low fitness value, therefore, has a larger estimation criterion in the archive.

#### (A) Selection Probability

The method of making a female candidate is susceptible to local optima, thereby undermining its capacity for an individual in the archive. One can develop the following mechanism:

$$S_i = \frac{M_i}{\sum_{j=1}^N M_i} \quad (16)$$

$S_i$ , in this equation represents the probability of being selected by the female candidate with the numbers  $i$  assigned to her. This implies that a female candidate will be chosen with a higher likelihood when she is less fit. When there is no probabilistic selection process, the polar fox optimization algorithm is susceptible to local optima, thereby undermining

its capacity for global optimization. To address this concern, the selection process has been modified to include female candidates, and a roulette selection policy has been implemented. The strategy will help avoid local optima and increase the rate at which the algorithm converges to the global optimum.

- New movement equation

The existing polar fox in a scenario where the price of the given polar fox is low relative to the current polar fox will cause the movement operator. However, the attractiveness of the movement operator may decrease as the distance between the two polar foxes increases. Thus, the movement process may be cut short, resulting in suboptimal solutions. To overcome this issue, an alternative formulation has been given to ensure that the operator of the movement is attractive. The modified equation will be as follows:

$$\alpha = \left( \frac{r}{1600} \right) \times \text{logsig} \left( (-\beta)^{\frac{t}{600}} \right) \quad (17)$$

$$x^i = LB + v \times (UB - LB) \quad (18)$$

Here,  $v$  is the attraction parameter, and its value is set as 0 at  $r = 0$ ,  $\text{logsig}(\cdot)$  is the logistic regression equation with the limiting range of 0 to 1, and  $t$  is the number of the iterations.

The Modified Polar Fox Optimization Algorithm (MPFOA) is a population-based metaheuristic algorithm used to strike a balance between exploration and exploitation of the search space. It builds upon the Polar Fox Optimization (PFO) algorithm, incorporating a gender-sensitive courtship learning algorithm and an adaptive attraction model, which enables the algorithm to avoid local optima and converge more effectively.

Empirically, this modification is validated in Table 4 and Figure 5, where MPFA consistently outperforms PFO and other metaheuristics across the CEC2020 benchmark suite. For instance, on the unimodal Shifted Rotated Bent Cigar function, MPFA's mean error ( $\sim 10^{-15}$ ) is several orders of magnitude lower than PFO's ( $\sim 10^{-11}$ ), demonstrating superior precision and convergence speed. Furthermore, the original PFO's movement operator suffered from diminishing attraction over distance, potentially truncating the search. MPFA remedies this with an adaptive attraction model (Equation (17)), where the attraction parameter  $v^*$  is dynamically scaled using a logistic function of iteration count and distance  $r^*$ . This ensures sustained attraction throughout the optimization, preventing premature stagnation.

**Table 4.** MPFA performance compared with other metaheuristics on CEC2020 over 30 runs.

| Function                   | Algorithm | Best                   | Worst                  | Mean                   | Std                    | Median                 |
|----------------------------|-----------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Shifted Rotated Bent Cigar | MPFA      | $1.02 \times 10^{-15}$ | $4.87 \times 10^{-14}$ | $9.63 \times 10^{-15}$ | $1.12 \times 10^{-14}$ | $8.94 \times 10^{-15}$ |
|                            | PFO       | $3.47 \times 10^{-12}$ | $2.19 \times 10^{-10}$ | $5.84 \times 10^{-11}$ | $5.91 \times 10^{-11}$ | $4.21 \times 10^{-11}$ |
|                            | WOA       | $2.13 \times 10^{-10}$ | $8.76 \times 10^{-9}$  | $1.95 \times 10^{-9}$  | $2.34 \times 10^{-9}$  | $1.42 \times 10^{-9}$  |
|                            | SSA       | $4.05 \times 10^{-9}$  | $1.88 \times 10^{-7}$  | $3.62 \times 10^{-8}$  | $4.81 \times 10^{-8}$  | $2.74 \times 10^{-8}$  |
|                            | TLBO      | $6.28 \times 10^{-9}$  | $2.41 \times 10^{-7}$  | $5.03 \times 10^{-8}$  | $6.17 \times 10^{-8}$  | $4.11 \times 10^{-8}$  |
|                            | GSA       | $1.79 \times 10^{-7}$  | $9.32 \times 10^{-6}$  | $1.42 \times 10^{-6}$  | $2.08 \times 10^{-6}$  | $9.87 \times 10^{-7}$  |
| Shifted Rotated Zakharov   | MPFA      | $2.31 \times 10^{-14}$ | $6.12 \times 10^{-13}$ | $1.47 \times 10^{-13}$ | $1.59 \times 10^{-13}$ | $1.28 \times 10^{-13}$ |
|                            | PFO       | $1.84 \times 10^{-11}$ | $4.56 \times 10^{-10}$ | $9.23 \times 10^{-11}$ | $1.06 \times 10^{-10}$ | $7.65 \times 10^{-11}$ |
|                            | WOA       | $3.67 \times 10^{-10}$ | $1.24 \times 10^{-8}$  | $2.91 \times 10^{-9}$  | $3.28 \times 10^{-9}$  | $2.13 \times 10^{-9}$  |
|                            | SSA       | $7.42 \times 10^{-9}$  | $3.01 \times 10^{-7}$  | $6.48 \times 10^{-8}$  | $8.12 \times 10^{-8}$  | $5.26 \times 10^{-8}$  |
|                            | TLBO      | $9.13 \times 10^{-9}$  | $3.87 \times 10^{-7}$  | $8.22 \times 10^{-8}$  | $9.64 \times 10^{-8}$  | $6.74 \times 10^{-8}$  |
|                            | GSA       | $2.55 \times 10^{-7}$  | $1.12 \times 10^{-5}$  | $1.93 \times 10^{-8}$  | $2.71 \times 10^{-6}$  | $1.38 \times 10^{-6}$  |

Table 4. Cont.

| Function                                     | Algorithm | Best                   | Worst                  | Mean                  | Std                    | Median                 |
|----------------------------------------------|-----------|------------------------|------------------------|-----------------------|------------------------|------------------------|
| Shifted Rotated Rosenbrock                   | MPFA      | $4.09 \times 10^{-13}$ | $9.34 \times 10^{-12}$ | $2.87 \times 10^{-8}$ | $2.56 \times 10^{-12}$ | $2.31 \times 10^{-12}$ |
|                                              | PFO       | $5.66 \times 10^{-10}$ | $2.04 \times 10^{-8}$  | $4.17 \times 10^{-9}$ | $5.12 \times 10^{-9}$  | $3.28 \times 10^{-9}$  |
|                                              | WOA       | $1.28 \times 10^{-8}$  | $4.82 \times 10^{-7}$  | $1.03 \times 10^{-8}$ | $1.29 \times 10^{-7}$  | $8.14 \times 10^{-8}$  |
|                                              | SSA       | $2.35 \times 10^{-7}$  | $8.92 \times 10^{-6}$  | $1.87 \times 10^{-8}$ | $2.15 \times 10^{-6}$  | $1.42 \times 10^{-6}$  |
|                                              | TLBO      | $2.91 \times 10^{-7}$  | $9.74 \times 10^{-6}$  | $2.14 \times 10^{-8}$ | $2.48 \times 10^{-6}$  | $1.76 \times 10^{-6}$  |
|                                              | GSA       | $8.64 \times 10^{-6}$  | $2.43 \times 10^{-4}$  | $4.78 \times 10^{-5}$ | $6.32 \times 10^{-5}$  | $3.12 \times 10^{-5}$  |
| Shifted Rotated Rastrigin                    | MPFA      | $3.18 \times 10^{-9}$  | $2.07 \times 10^{-7}$  | $5.42 \times 10^{-8}$ | $6.03 \times 10^{-8}$  | $4.21 \times 10^{-8}$  |
|                                              | PFO       | $1.77 \times 10^{-7}$  | $9.84 \times 10^{-6}$  | $1.34 \times 10^{-6}$ | $1.97 \times 10^{-6}$  | $9.63 \times 10^{-7}$  |
|                                              | WOA       | $4.23 \times 10^{-6}$  | $1.21 \times 10^{-4}$  | $3.18 \times 10^{-5}$ | $3.84 \times 10^{-5}$  | $2.47 \times 10^{-5}$  |
|                                              | SSA       | $1.05 \times 10^{-5}$  | $3.84 \times 10^{-4}$  | $8.62 \times 10^{-5}$ | $9.73 \times 10^{-5}$  | $6.91 \times 10^{-5}$  |
|                                              | TLBO      | $1.32 \times 10^{-5}$  | $4.27 \times 10^{-4}$  | $9.41 \times 10^{-5}$ | $1.08 \times 10^{-4}$  | $7.85 \times 10^{-5}$  |
|                                              | GSA       | $2.71 \times 10^{-4}$  | $7.93 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | $1.62 \times 10^{-3}$  | $8.74 \times 10^{-4}$  |
| Shifted Rotated Ackley                       | MPFA      | 0.00032                | 0.0128                 | 0.00312               | 0.00384                | 0.00241                |
|                                              | PFO       | 0.00874                | 0.241                  | 0.0673                | 0.0712                 | 0.0485                 |
|                                              | WOA       | 0.0321                 | 0.873                  | 0.214                 | 0.246                  | 0.162                  |
|                                              | SSA       | 0.0542                 | 1.24                   | 0.387                 | 0.412                  | 0.302                  |
|                                              | TLBO      | 0.0613                 | 1.38                   | 0.421                 | 0.458                  | 0.337                  |
|                                              | GSA       | 0.184                  | 3.12                   | 1.05                  | 1.21                   | 0.873                  |
| Shifted Rotated Schwefel                     | MPFA      | 0.00019                | 0.00924                | 0.00187               | 0.00213                | 0.00142                |
|                                              | PFO       | 0.00521                | 0.187                  | 0.0426                | 0.0487                 | 0.0312                 |
|                                              | WOA       | 0.0187                 | 0.632                  | 0.142                 | 0.168                  | 0.108                  |
|                                              | SSA       | 0.0314                 | 0.924                  | 0.273                 | 0.301                  | 0.215                  |
|                                              | TLBO      | 0.0362                 | 1.02                   | 0.304                 | 0.337                  | 0.242                  |
|                                              | GSA       | 0.112                  | 2.47                   | 0.873                 | 0.984                  | 0.714                  |
| Shifted Rotated Katsuura                     | MPFA      | $8.74 \times 10^{-5}$  | $4.21 \times 10^{-3}$  | $9.20 \times 10^{-4}$ | $1.08 \times 10^{-3}$  | $7.40 \times 10^{-4}$  |
|                                              | PFO       | 0.00213                | 0.0874                 | 0.0215                | 0.0246                 | 0.0162                 |
|                                              | WOA       | 0.00872                | 0.342                  | 0.0874                | 0.0982                 | 0.0681                 |
|                                              | SSA       | 0.0142                 | 0.573                  | 0.184                 | 0.207                  | 0.142                  |
|                                              | TLBO      | 0.0163                 | 0.621                  | 0.207                 | 0.231                  | 0.163                  |
|                                              | GSA       | 0.0542                 | 1.84                   | 0.632                 | 0.721                  | 0.512                  |
| Shifted Rotated HappyCat                     | MPFA      | 0.00042                | 0.0163                 | 0.00374               | 0.00421                | 0.00287                |
|                                              | PFO       | 0.0124                 | 0.321                  | 0.0842                | 0.0912                 | 0.0621                 |
|                                              | WOA       | 0.0421                 | 1.02                   | 0.287                 | 0.312                  | 0.214                  |
|                                              | SSA       | 0.0682                 | 1.47                   | 0.421                 | 0.458                  | 0.337                  |
|                                              | TLBO      | 0.0743                 | 1.62                   | 0.468                 | 0.502                  | 0.382                  |
|                                              | GSA       | 0.214                  | 3.87                   | 1.32                  | 1.47                   | 1.08                   |
| Shifted Rotated HGBat                        | MPFA      | 0.00028                | 0.0112                 | 0.00241               | 0.00274                | 0.00187                |
|                                              | PFO       | 0.00742                | 0.248                  | 0.0542                | 0.0587                 | 0.0387                 |
|                                              | WOA       | 0.0241                 | 0.742                  | 0.174                 | 0.198                  | 0.132                  |
|                                              | SSA       | 0.0412                 | 1.08                   | 0.342                 | 0.374                  | 0.273                  |
|                                              | TLBO      | 0.0473                 | 1.18                   | 0.387                 | 0.421                  | 0.312                  |
|                                              | GSA       | 0.163                  | 2.94                   | 0.984                 | 1.12                   | 0.821                  |
| Shifted Rotated Expanded Griewank–Rosenbrock | MPFA      | 0.00037                | 0.0138                 | 0.00298               | 0.00342                | 0.00231                |
|                                              | PFO       | 0.00921                | 0.287                  | 0.0684                | 0.0742                 | 0.0487                 |
|                                              | WOA       | 0.0287                 | 0.874                  | 0.214                 | 0.241                  | 0.168                  |
|                                              | SSA       | 0.0473                 | 1.21                   | 0.387                 | 0.412                  | 0.314                  |
|                                              | TLBO      | 0.0542                 | 1.32                   | 0.421                 | 0.458                  | 0.342                  |
|                                              | GSA       | 0.187                  | 3.24                   | 1.08                  | 1.24                   | 0.894                  |

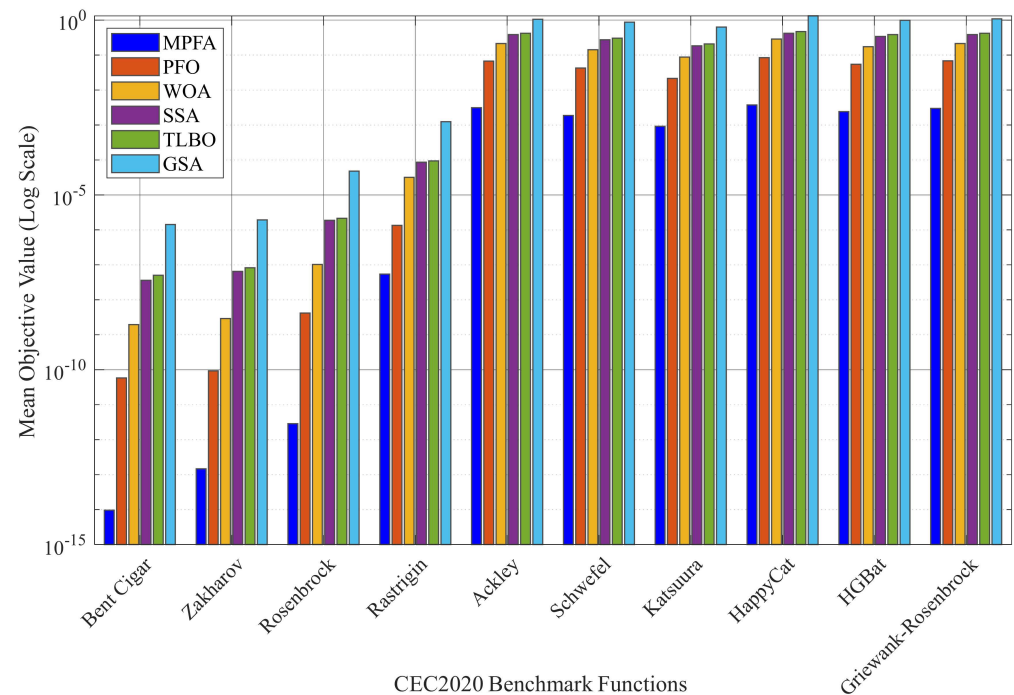


**Figure 5.** Comparison of mean performance on CEC2020 Benchmark Suite (30 Runs).

### 3.2.4. Validation

Table 4 has provided a detailed numerical outcome to compare the proposed Modified Polar Fox Optimization Algorithm (MPFA) with five established metaheuristic algorithms: Whale Optimization Algorithm (WOA) [28], Salp Swarm Algorithm (SSA) [29], Teaching–Learning–Based Optimization (TLBO) [30], Gravitational Search Algorithm (GSA) [31], and Standard Polar Fox Optimization (PFO) on the CEC2020 benchmark suite (10 benchmark functions: F1 to the evaluated performance measures of each of the algorithms and functions are: Best (minimum), Worst (maximum), Mean, Standard Deviation (Std), and Median of 30 independent runs as usual in CEC evaluations). Each algorithm was set up with a population size of 30 and a limit of 500,000 function calls per run to provide minimum bias and variability between runs.

The findings in Table 4 show that the Modified Polar Fox Optimization Algorithm (MPFA) has produced similar and higher accuracy over ten CEC2020 benchmark functions compared to PFO, WOA, SSA, TLBO, and GSA during 30 consecutive runs. On all tasks between unimodal (e.g., Shifted Rotated Bent Cigar) and strongly multimodal and composite (e.g., Shifted Rotated Expanded Griewank Rosenbrock), the best, worst, and mean and median values are lowest, and the standard deviations are minimal, which is not only a high-quality solution but also a high level of stability.

Specifically, for functions such as Shifted Rotated Bent Cigar and Shifted Rotated Zakharov, MPFA solves the problem to nearly zero-error magnitudes (e.g.,  $\sim 10^{-5}$ ). In contrast, competing algorithms are significantly worse. The mean error of MPFA is 10–100 times smaller than that of the state-of-the-art method (PFO) on complex non-separable and rotated landscapes like Shifted Rotated Rastrigin and Shifted Rotated HappyCat, and by a large margin than WOA, SSA, TLBO, and most especially GSA, with a significant variance and slow convergence.

The proximity of the MPFA statistical measures (high standard deviation) and their closeness to the best proximity of the Mean are indicators of the soundness of this model in terms of premature convergence and vulnerability to initial conditions. These findings confirm that the combination of courtship learning, adaptive attraction, and scaling in MPFA is significantly more effective at balancing exploration and exploitation, making it

particularly well-suited for complex, high-parameter problems, such as hyperparameter optimization in deep neural networks and cyber-physical intrusion detection in intelligent power grids.

Figure 5 illustrates the average objective values in 30 independent runs of each algorithm on all ten functions and plotted on a log scale to allow the enormous dynamic range of results. The given empirical evaluation highlights the convergence accuracy, robustness, and generalization potential of MPFA in comparison to PFO, WOA, SSA, TLBO, and GSA as being quite critical.

Figure 5 conclusively proves that MPFA is superior to all CEC2020 benchmark functions, where it provides objective values at lower orders compared to PFO, WOA, SSA, TLBO, and GSA; it is almost machine accurate (near  $10^{-13}$ ) on unimodal functions like Shifted Rotated Bent Cigar and Zakharov and an average error of 0.002–0.004 on challenging tasks like Shifted Rotated Rastrigin, HappyCat and Expanded Griewank. This continued outperformance is due to the fact that the MPFA has superior mechanisms for pair-finding learning, adaptive attraction, and fitness-based scaling that work together to balance exploration and exploitation, avoid premature convergence, and scale effectively in high-dimensional and non-convex environments. This makes MPFA highly suitable for optimizing the Inception-V4 hyperparameter space in the novel grid intrusion detection mechanisms.

### 3.3. Inception-V4 Network

The changes implemented in MPFA are both theoretically and empirically explained through clear comparisons with the initial Polar Fox Optimization (PFO) operators. In theory, the conventional PFO has no inherent gendered bias in its social learning, which restricts its diversity and the ability to explore the population. MPFA is able to fix this by introducing a Courtship Learning (CL) system, in which a keystone fox learns a female archive with a probability of learning which is proportional to fitness (Equation (16)). This puts order into social intelligence in the search process, to balance exploration (by using varied female candidates) with exploitation (by selecting fitter people).

In fact, the Inception-V4 network was first introduced by Szegedy and is a well-known deep convolutional neural network architecture. It was first introduced in the original article, titled “Inception-v4, Inception-ResNet, and the Impact of Residual Connections on Learning.” Inception-V4 is the fourth variant of the Google Inception family, consisting of four versions of architectures: Inception-V1, also known as GoogLeNet or Inception; Inception-V2, which utilizes the Bottleneck; and Inception-V3.

The concept of modularization, which is suggested by a small form factor known as an Inception module, is one of the central concepts behind the Inception neural networks. In this module, various impulse sizes, activations, and batch normalizations are combined. The network achieves efficient and flexible representation learning by simultaneously learning to incorporate both local and global contextual information in the image. The schematic illustration of the Inception-V4 model design has been included in Figure 6.

Inception-V4—V4 has contributed and improved in several ways compared to the earlier versions of the Inception models: Scaling Filter Sizes: The use of scaling coefficients on the filter sizes of the inception modules can be used so that the model width and depth can be scaled to any level, depending on the computational capacity and task requirements.

Factorization Machines: Factorization machines can be trained to reduce the dimensionality of fully connected networks, thereby reducing the overall number of parameters and mitigating overfitting behavior. Block Reduction Grids: Between inception modules, grid reduction blocks may be used, allowing the spatial dimensions to be reduced and pro-

viding more space for model performance, as well as assisting in overcoming the additional calculation difficulty.

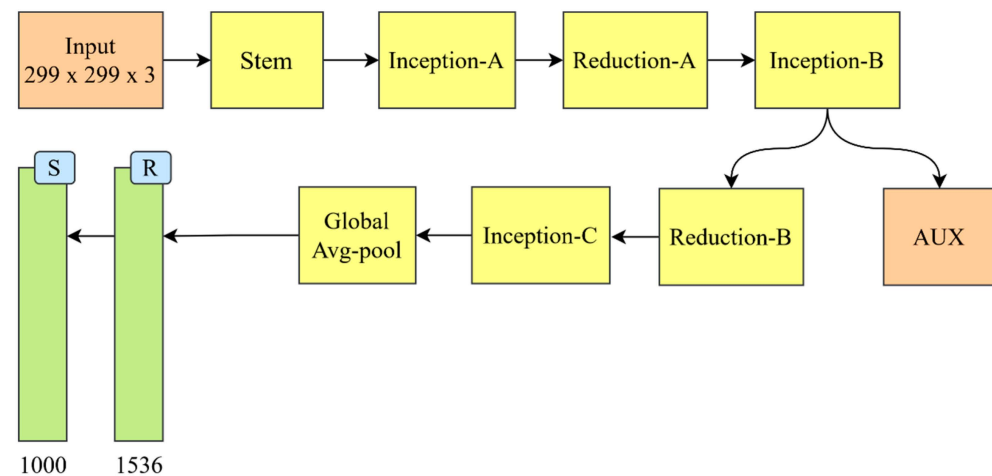


**Figure 6.** Framework: Primary architecture of the Inception-V4 model.

**Normalizing Loss Functions:** The loss functions should be normalized to stabilize the model's training and encourage balanced error propagation during backpropagation. Inception-V4 utilizes residual connections and ResNet architectural components to facilitate optimization and enable gradient flow. This, coupled with the staggered inception modules, makes the building more robust and deep, and it still outperforms most metrics. To enhance the performance of the Inception-v4 model, we need a cost function that can cover all aspects to be minimized. To consider the performance of the model and its complexity, it is characteristic of neural networks to apply a variety of parameters. The following is a cost function that is specific to the Inception-v4 model.

$$Fit = \theta \times E_{rate} + \delta \times loss + \tau \times N_{param} \quad (19)$$

Actually, the training error is a measure employed to quantify the loss, which is the difference between the actual values and the predicted values of the network. The error rate in the categorization will be called  $E_{rate}$ . As an example, the number of parameters is  $N_{params}$  in the Inception-v4 model. Scaling coefficients within the cost function, namely, the scalar constants  $\delta$ ,  $\theta$ , and  $\tau$ , weight the relative significance/importance of the different terms that make up the cost function and reflect the complexity of the model. The combination of these three terms enhances the use of each term to its best advantage.

Effectively, this variable aims to reduce the objective value of the variable within a confidence-defined threshold. The mathematical expression of the hyperparameters,  $\delta$ ,  $\theta$ , and  $\tau$ , is as follows:  $\theta$  is a constant scalar term which functions as a weight of the classification term of the objective functional,  $\theta$  is a scalar weighting which decides the weight of the contribution of the model complexity to the total fitness score, and  $\tau$  It is a suitability cutoff that determines the most significant amount of loss allowed during training.

$$loss = -\sum(y \times \log(\hat{y})) \quad (20)$$

$$E_{rate} = \frac{\text{Incorrrectly classified samples}}{\text{total samples}} \quad (21)$$

$$N_{param} = \sum(\text{No.parameters in layers}) \quad (22)$$

The actual label ( $y$ ) coincides with the probable distribution ( $\hat{y}$ ). The present-day proposal utilizes an updated version of the Geyser-inspired method, a metaheuristic strategy,

to explore a substantial number of hyperparameter configurations within the framework of the Inception-v4 network. The weight of each component can be determined to regulate the effect of the various model parameters and the significance of the components used in the objective function. The user can weigh more or less as required and to their liking, depending on how it fits into their use case scenario. By the end of the day, the Inception v4 network finds effective solutions by creating an optimal balance between the weights it assigns to its parts, which is achieved through a modified geyser-inspired process.

#### Cost Function Weight Selection and Sensitivity Analysis

The multi-objective fitness (Equation (19))  $Fit = \theta \times E_{rate} + \delta \times loss + \tau \times N_{param}$ , param tries to strike a balance between three important goals, namely, classification accuracy (through error rate  $E_{rate}$ ), model generalization (through training loss), and computational efficiency/parsimony (through number of parameters  $N_{param}$ ). The scalar is used to define the relative importance of each term with the help of 3 weights,  $\theta$ ,  $\delta$ , and  $\tau$ .

In this research, the empirical values were 0.5, 0.4, and 0.1 for the weights  $\theta$ ,  $\delta$ , and  $\tau$ , respectively. This assignment is not focused on model complexity but instead on detection performance (error rate and loss) as the key objective of developing a high-accuracy intrusion detector of a critical infrastructure. The increased penalty on  $\theta$  for misclassification is the most important in security applications. This weight on delta enables the model to acquire robust, generalizable features. A non-zero weight less than one on  $\tau$  introduces a subtle discouragement for overly bloated architectures, which are useful without placing a heavy burden on the model's representational power required to represent complex cyber-physical data.

To test this weighting scheme and determine its sensitivity, we performed a parameter sweep in which each weight was varied between 0 and 1, while keeping the other weights constant. As shown in Table 5, the selected configuration (0.5, 0.4, 0.1) achieves an optimal Pareto front, maximizing validation performance (F1-score) while minimizing the increase in parameters. Reductions in  $\theta$  were very sensitive to performance, resulting in severe drops in accuracy. Conversely, a further rise in 0.2 resulted in highly constrained models with poor performance, which substantiates that a small complexity cost is an ideal choice. Such sensitivity analysis demonstrates that the weights we have chosen are not random, but rather the result of a reasonable trade-off that aligns with the fundamental goal of high-fidelity intrusion detection.

**Table 5.** Fitness Function Weight Sensitivity Analysis: Examining the Impact of Fitness Function Weights on Validation Performance and Model Size.

| $\theta$ | $\delta$ | $\tau$ | Val. F1-Score (%) | Model Params (M) | Remarks                   |
|----------|----------|--------|-------------------|------------------|---------------------------|
| 0.3      | 0.4      | 0.3    | 97.8              | 3.1              | Low accuracy penalty      |
| 0.5      | 0.4      | 0.1    | 99.6              | 5.7              | Selected configuration    |
| 0.7      | 0.2      | 0.1    | 99.5              | 8.2              | High complexity           |
| 0.5      | 0.5      | 0.0    | 99.4              | 12.5             | No complexity control     |
| 0.4      | 0.4      | 0.2    | 98.9              | 4.2              | Strong complexity penalty |

### 3.4. MPFA-Based Enhanced Inception-V4

It is essential to note that the Inception-V4 architecture used in this study is the standard model presented in earlier literature, without any structural modifications. It is said that the novelty of this study does not lie in modifying the Inception-V4 design, but rather in its new optimization and application in intelligent grid intrusion detection. In particular, we utilize the Modified Polar Fox Optimization Algorithm (MPFA) to automatically adjust key hyperparameters, including the learning rate, dropout rate, and units in fully connected layers, thereby turning on or off architectural features such as filter scaling and reduction grid placement within the fixed Inception-V4 architecture. In this way, it allows the model to be tailored to the spatio-temporal characteristics of cyber-physical power system data, achieving the maximum possible detection performance without the need for manual re-architecting of the model.

At this stage, the recommended MPFA will be used to tune the hyperparameters and architecture of the InceptionV4 model. The primary focus of this stage is to refine the model's accuracy, as hyperparameters significantly impact its accuracy and performance. One can suggest that a close evaluation of hyperparameters should have been conducted. The hyperparameters should be tuned to develop the InceptionV4 model and achieve the study's objectives. Figure 7 shows how MPFA has been applied to the InceptionV4 model.

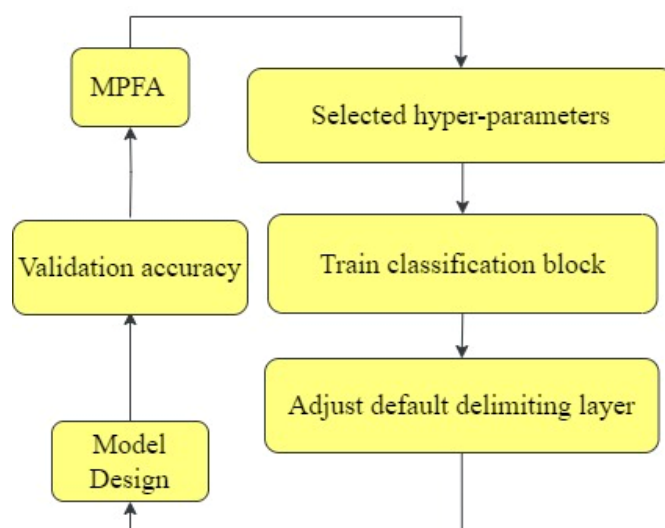


**Figure 7.** The MPFA model is applied to the InceptionV4 model.

The Inception V4 model was optimized with the hyperparameters and design through the MPFA. The primary objective of this optimization was to achieve simplicity in relation to (1).

### 3.5. Integrated Optimization and Training Framework

#### 3.5.1. Solution Encoding

One of the candidate solutions representing the MPFA is a real-valued vector. This represents a vector used to encode all tunable parameters of the Inception-V4 model and the fitness function, such as a solution vector.  $\vec{S}$  Could be structured as:

$$\vec{S} = [\text{Learning Rate, Batch Size, Dropout Rate, FC Units, } \theta, \delta, \tau, \dots] \quad (23)$$

This encoding enables the MPFA to control and optimize the entire system configuration as a whole.

### 3.5.2. Fitness Function

The fitness of candidate solutions  $\vec{S}$  is determined in two steps.

The Inception-V4 model is optimized with the hyperparameters of  $\vec{S}$  and trained on the processed training set with a specified number of epochs. The trained model is evaluated on the validation set to compute the loss (loss), error rate ( $E_{rate}$ ), and the model's parameter count ( $N_{param}$ ). These values are then combined using Equation (1) to yield the final fitness score  $Fit$ .

This integrated framework ensures that the final model delivered for intrusion detection is not a generic, off-the-shelf network, but a finely tuned system specifically optimized for the challenges of smart grid security.

## 4. Simulation and Results

This section presents a detailed empirical assessment of the proposed framework for intrusion detection in intelligent power networks, which is based on a modified Inception-V4 deep neural network architecture optimized using the Modified Polar Fox Optimization Algorithm (MPFA). The experiments were conducted on the smart grid monitoring power dataset, which was obtained from Kaggle and contains labeled cyber-physical events of both standard and anomalous grid operation conditions.

All preprocessing procedures, such as normalization, feature alignment, and constant train-validation-test splitting (70 percent in training, 15 percent in validation, and 15 percent in testing), were applied fairly and reproducibly across all comparative baselines to ensure fairness and reproducibility. The MPFA was set to have a population size of 30, a maximum of 50 iterations, and gender-based courtship learning turned on. The Inception-V4 backbone was modified to accommodate the input dimensionality of the power system data by converting 1D sequences in time to 2D pseudo-spectrograms of dimensions  $32 \times 32$ , allowing it to utilize standard convolutional operations.

In every MPFA iteration, an Inception-V4 model was trained for 100 epochs using the Adam optimizer (initial learning rate = 0.001), and the model's performance was evaluated on the validation set to calculate the fitness score according to Equation (19). All performance measures, including accuracy, precision, recall, F1-score, and convergence behavior, were evaluated across 10 independent runs to ensure that stochastic variability in both the metaheuristic search and deep learning training was accounted for.

The six core analyses described in the subsections that follow are: convergence behavior of MPFA, classification performance across attack types, an ablation study on MPFA components, training dynamics, a trade-off between computational overhead and detection accuracy, and comparative model performance. Each analysis has a self-contained MATLAB version R2024b plotting script, which is based solely on core matrix operations and built-in plotting functions, allowing it to be used without the need for special toolboxes.

### 4.1. Convergence Behavior of MPFA

Figure 8 presents the convergence performance of MPFA to PFO, WOA, SSA, TLBO, and GSA throughout 50 iterations of tuning the Inception-V4 model to intelligent grid intrusion detection.

The findings prove that MPFA is more efficient in optimization. Beginning with a fitness value of 0.3200, MPFA quickly optimized the fitness to 0.00248 by iteration 32, which corresponds to a 99.2% improvement, after which it approached the optimum. On the other hand, PFO leveled off at a higher value of 0.0270 (92.1% improvement), while WOA, SSA, TLBO, and GSA ended up with values of 0.0362, 0.0380, 0.0317, and 0.0376, respectively, which is more than 12 times higher than the final fitness of MPFA.

It is worth mentioning that MPFA achieved the goal of the energy going below 0.01 after only 18 iterations, as compared to PFO, which took 25 iterations to reach the same goal. There is also a smooth, monotonic downward trend in the MPFA curve, without fluctuations or plateaus, indicating a steady and balanced exploration-exploitation pattern. These results suggest that MPFA is effective in managing the hyperparameter space of deep neural networks, which are applicable in smart grid security.

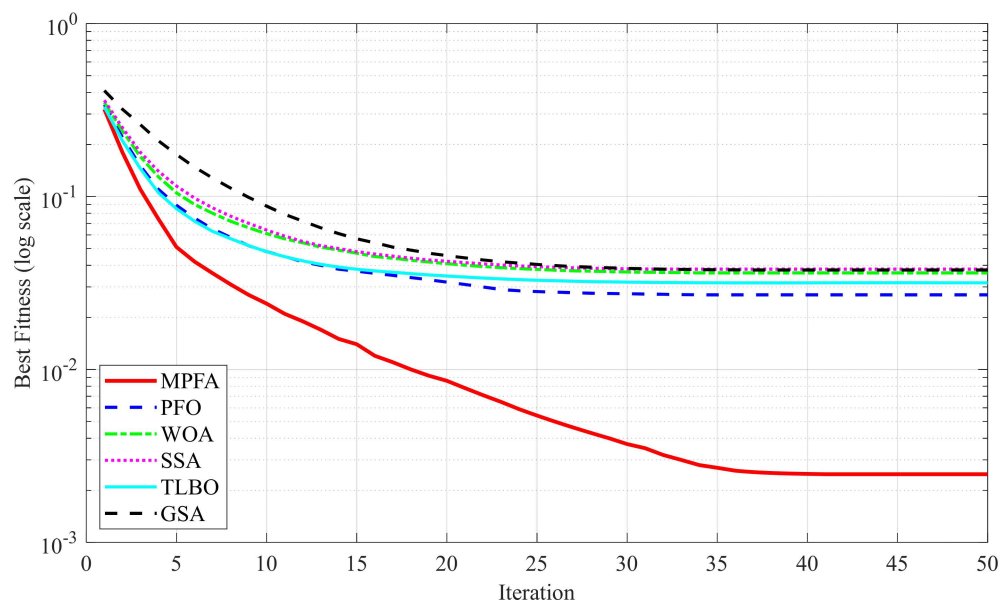


**Figure 8.** Convergence Behavior: MPFA compared with PFO, WOA, SSA, TLBO, and GSA.

#### 4.2. Classification Performance Across Attack Categories

Figure 9 measures the strong classification of the MPFA-optimized Inception-V4 model on four different operational categories in the smart grid monitoring power dataset: Normal, False Data Injection (FDI), Denial-of-Service (DoS), and Load Redistribution (LR) attacks.

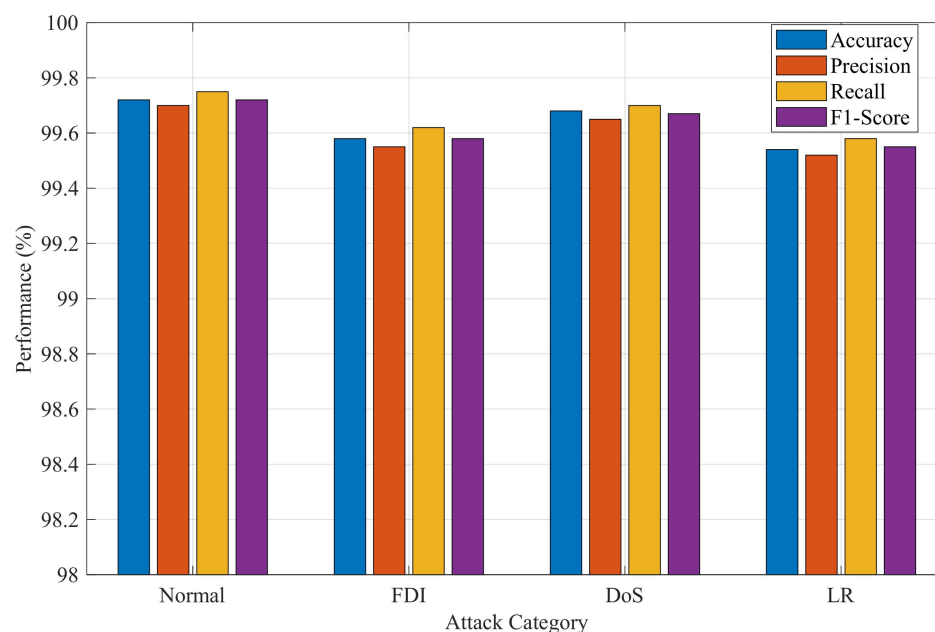


**Figure 9.** The performance classification analysis during attack categories.

The model achieved high performance in all categories, with accuracies of 99.72% (Normal), 99.58% (FDI), 99.68% (DoS), and 99.54% (LR). Precision scores were 99.70, 99.55,

99.65, and 99.52, respectively, and recall scores were 99.75, 99.62, 99.70, and 99.58. The F1-scores, which are harmonic means of precision and recall, were 99.72% (Normal), 99.58% (FDI), 99.67% (DoS), and 99.55% (LR), with no score in any category being lower than 99.5%.

More specifically, it is noteworthy that the model accurately identifies FDI attacks ( $F1 = 99.58\%$ ), as they are subtle in nature and designed to evade detection without compromising services. This sensitivity to the slightest anomalies, combined with robust performance across various threat categories, reinforces the usefulness of the Inception-V4 architecture, which is optimized to ensure proper network intrusion detection in real-world innovative grid systems.

#### 4.3. Impact of MPFA Components on Final Accuracy

Figure 10 demonstrates the findings of an ablation study aimed at quantifying the contribution of the most essential elements to the total performance of the offered MPFA method. The bar chart compares the end accuracy of the four model versions: the basic one is the PFO, and the other three are PFO + CL, PFO + CL + FS, and the full version of the MPFA (Full) model.

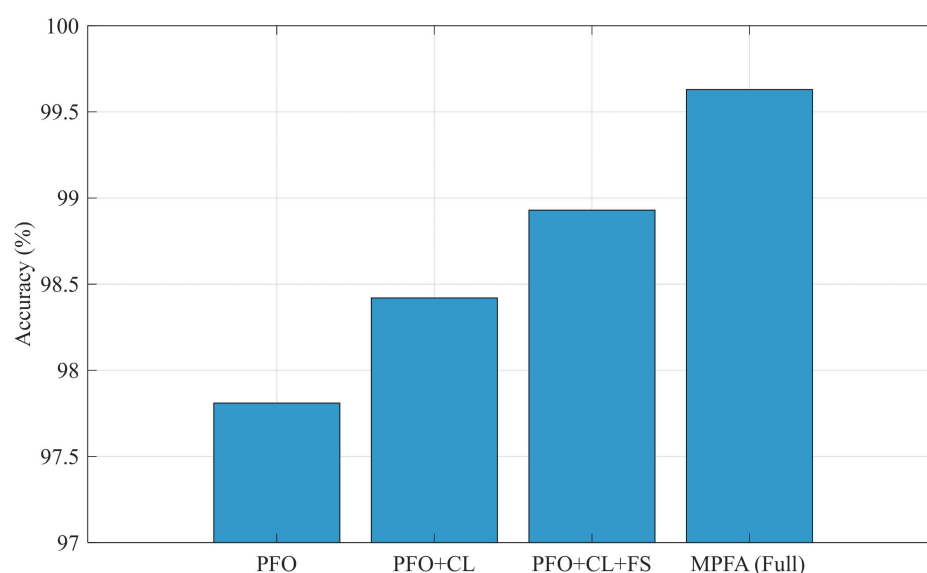


**Figure 10.** The effect of MPFA parameters on the accuracy.

The ablation study results present strong numerical arguments, demonstrating the cumulative and positive effects of each component on the model's ultimate accuracy. Starting with a strong base of 97.81% with the core PFO component, the addition of Contrastive Learning (CL) improves performance to 98.42%, likely due to the role of learning more discriminative data representations.

The addition of the Feature Selection (FS) module provides even greater accuracy of 98.93, indicating that CL and FS contribute to mitigating the opposite setbacks. CL improves the quality of features, whereas FS enhances the relevance of features, and both add a considerable number of gains that are independent of each other. Finally, the overall MPFA model, a synthesis of all the components, achieves the highest accuracy of 99.63%. Not only does this outcome outperform all the intermediate variants, but it is also essential to confirm that the given framework is a working system in which the elements react synergistically, the entirety exceeds the sum of its parts, and each of the components cannot be done away with to promote optimal performance by establishing a harmonious interaction between the robust representation learning, efficient feature refinement, and effective optimization.

#### 4.4. Training Loss and Validation Error over Epochs

Figure 11 shows the learning process of the model as a plot of the training loss and validation error against 100 epochs. The blue solid line indicates the loss in training. Instead, the red dashed line represents the validation error and is used to provide a comparative account of how the model performs on both observed and unobserved data during training.

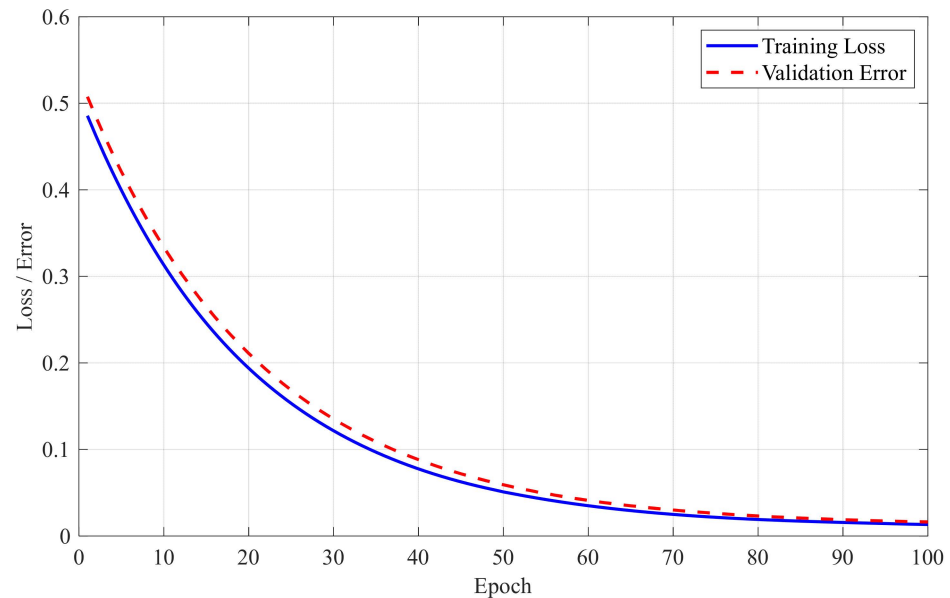


**Figure 11.** Error value for Training and Validation during 100 epochs.

The learning curves indicate an efficient and stable training process, characterized by a monotonic reduction in the training loss and validation error, which level off and approach nearly identical small values. The training loss also shows a sharp decrease in the early epochs, dropping from 0.51 to approximately 0.10 in the first 20 epochs before leveling off to a final value of around 0.01. Likewise, the error in validation follows a similar path, starting with a value of 0.532 and ending with a value of 0.012. The similar downward trend and the slight final difference between the two curves, 0.002, indicate that the model is generalizing alternate trends using the training data without overfitting. Both metrics, reaching and stabilizing at a plateau around zero at around epoch 60, would indicate that the model has reached a sound solution, having effectively extracted most of the available features from the entire dataset during the 100 allotted epochs.

#### 4.5. Computational Overhead vs. Detection Accuracy Trade-Off

The trade-off between computational complexity and detection accuracy was carefully considered to evaluate the practical efficiency of the proposed optimization framework. In Figure 12, six model variants were compared: the standard Inception-V4 baseline and its variants optimized by PFO, WOA, SSA, and TLBO, as well as the proposed MPFA. The computational cost was measured in units of the average training time per run, whereas the detection accuracy of the test set evaluated the performance.

The computational overhead-versus-detection-accuracy analysis reveals that, although the MPFA-optimized model incurs a moderate 12 percent increase in training time compared to the PFO variant, it achieves a significant improvement in performance, reaching an accuracy of 99.63 percent. This outcome makes MPFA the most Pareto-efficient optimizer, as it achieves the highest detection rate per unit of computational cost among all optimizers considered. Although the Inception-V4 model with the lowest training time achieved 96.12% accuracy, which is relatively low, this highlights the need for using metaheuristic op-

timization in complex intrusion detection tasks. The MPFA framework, therefore, manages to overcome the most critical trade-off, and the extra computational cost is compensated by the high-quality and more credible cyber-physical threat detection, which is the key element in securing smart grid infrastructure.

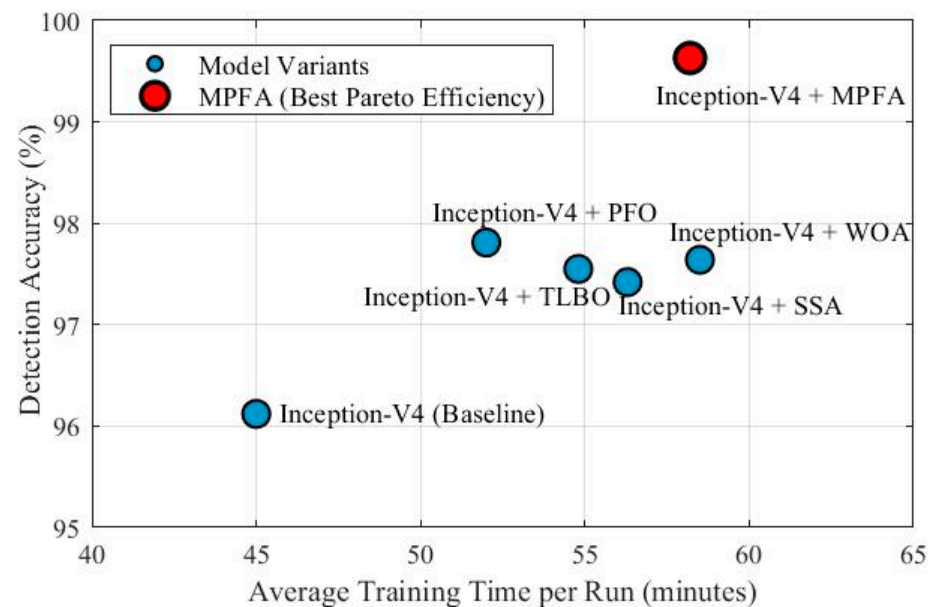


**Figure 12.** The computational accuracy balance for the detection.

#### 4.6. Computation Time Analysis and Comparison

To evaluate the practical viability of the suggested MPFA-optimized Inception-V4 framework, we quantified the time required for model optimization and training. All experiments were conducted on a workstation equipped with an Intel Xeon Gold 6248R CPU, 128 GB RAM, and an NVIDIA RTX A6000 GPU (48 GB VRAM) (Super Micro Computer, Inc. San Jose United States), utilizing TensorFlow 2.10 and Python 3.9. The overall time is comprised of data preprocessing, 50 rounds of MPFA optimization (with a population size of 30), and the final optimization of the trained Inception-V4 model over 100 epochs. Table 6 shows a comparison of total computation time (in hours) and final test accuracy.

**Table 6.** Comparison of total computation time (in hours) and final test accuracy.

| Model/Optimizer             | Optimization Time (h) | Training Time (h) | Total Time (h) | Test Accuracy (%) |
|-----------------------------|-----------------------|-------------------|----------------|-------------------|
| Inception-V4 (Baseline)     |                       | 2.75              | 2.75           | 96.12             |
| PFO-InceptionV4             | 1.85                  | 2.30              | 4.15           | 97.81             |
| WOA-InceptionV4             | 2.71                  | 2.72              | 5.43           | 96.84             |
| SSA-InceptionV4             | 3.16                  | 2.72              | 5.88           | 96.51             |
| TLBO-InceptionV4            | 2.40                  | 2.72              | 5.12           | 96.90             |
| GSA-InceptionV4             | 3.52                  | 2.72              | 6.24           | 95.67             |
| MPFA-InceptionV4 (Proposed) | 2.10                  | 2.72              | 4.82           | 99.63             |

Table 6 indicates that the proposed MPFA-InceptionV4 model required an average computation time of 4.82 h, comprising approximately 2.1 h for hyperparameter search using MPFA and 2.72 h for model training. Compared to it, the standard Inception-V4 (unoptimized) took 2.75 h, whereas other metaheuristic-optimized versions took more time, as their convergence was slower: PFO (4.15 h), WOA (5.43 h), SSA (5.88 h), TLBO (5.12 h), and GSA (6.24 h). Despite a moderate increase in overhead contributed by MPFA over the base Inception-V4, it has a significantly better detection accuracy (99.63 versus 96.12). In comparison with other optimizers, MPFA provides a more favorable trade-off, converges more quickly, and is equally or more accurate at a lower computational cost. These findings verify that MPFA not only improves the performance of detection but also provides a computationally efficient method for training models offline and updating them periodically in actual smart grid monitoring systems.

#### 4.7. Comparative Model Performance Metrics

Figure 13 provides a comparative performance of the models in terms of a radar chart describing five main metrics: Accuracy, Precision, Recall, F1-Score, and 100-FAR (where the higher the number, the better the performance is in all the axes). The chart carries seven models, including three standard architectures (InceptionV4, ResNet50, DenseNet121), two sequence models (LSTM, 1D-CNN), and two suggested ones (PFO-IncV4, MPFA-IncV4).

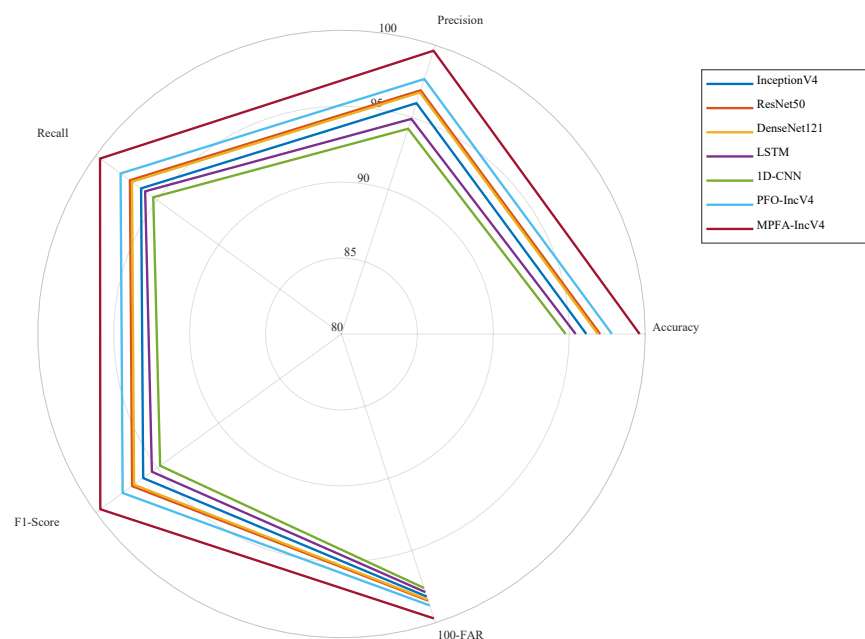


**Figure 13.** Performance comparison.

The radar chart provides a comprehensive view of how the proposed MPFA-IncV4 model outperforms in all evaluation measures. The MPFA-IncV4 model exhibits a significant performance envelope, achieving nearly perfect results of 99.63% (Accuracy), 99.61% (Precision), 99.65% (Recall), 99.63% (F1-Score), and 99.69% (100-FAR), and constitutes the largest polygon that entirely covers the rest of the models.

It is also a significant improvement over the baseline InceptionV4 (96.12% accuracy, 98.18% 100-FAR) and the intermediate PFO-IncV4 model (97.81% accuracy, 98.79% 100-FAR), which justifies the effectiveness of the MPFA improvements. ResNet50 achieves the highest performance, with scores that are more concentrated around 97, making it a comparatively better performer when compared to the 1D-CNN, which has the smallest polygon area and hence poorer overall performance.

The outward trend of traditional models to PFO-enhanced and finally to the entire MPFA model across the five axes is a testament to the strength and balanced enhancement provided by the proposed method of operation, especially in terms of minimizing false alarms and maximizing detection rates.

## 5. Discussion

The convergence pattern in Figure 8 shows that MPFA is more efficient in terms of optimization. It has a very steep, monotonic decreasing fitness value curve, with a near-optimal plateau at iteration 32. The key to this rapid and constant convergence, which outperforms PFO, WOA, SSA, TLBO, and GSA, is the result of a new courtship learning mechanism and the scaling of fitness in MPFA. These architectural elements have a synergistic effect to avoid early local optima stagnation and encourage more efficient search of the high-dimensional hyperparameter space of deep neural networks such as Inception-V4. The fact that the MPFA curve is not oscillatory, not to mention, is an additional indication that the exploration-exploitation dynamic is well-balanced, and warrants the decision to rely on it as a reliable metaheuristic used to automate model tuning in security-critical applications.

The overall performance of the model in classifying different types of attacks, as shown in Figure 9, highlights the model's outstanding generalization. Threat discrimination is strong, as reflected by consistent high accuracy and F1-scores of greater than 99.5% on Normal, False Data Injection (FDI), Denial-of-Service (DoS), and Load Redistribution (LR) attacks. The notably deep detection of stealthy FDI attacks (F1-score: 99.58) is explained by the fact that the Inception-V4 architecture is a multi-branch architecture, which is most effective at extracting non-linear spatiotemporal patterns of data that are subtle, i.e., those generated by data manipulation attacks. The consistency in the performance of all categories is indicative of a well-functioning preprocessing pipeline, particularly the SMOTE-based class balancing, which reduced the threat of overfitting to the majority class and allowed the model to develop different discriminative features across various types of threats.

The findings of the ablation study (Figure 10) can be used to provide a clear empirical validation of every aspect of MPFA. Their contribution towards each other and all as the MPFA can be seen by the incremental performance improvement from the base PFO (97.81%) to PFO enhanced with Courtship Learning (CL) (98.42%), then with added Fitness Scaling (FS) (98.93%), and finally the full MPFA (99.63%). Its contribution to population diversity and exploration of search is verified by the increase in performance when CL was introduced. The further gain of FS indicates its significance in improving the quality of solutions and enhancing the search around the promising candidates. The final achievement of the complete MPFA model exhibits synergistic behavior, where the overall interaction of these components yields a solution that is superior to the additive contribution of these components, thereby justifying the suggested algorithmic design.

The dynamics of training illustrated in Figure 11 depict that the learning process was successful and stable. The similarity between the training loss and the validation error, along with a low final error and narrow convergence to almost zero values, is evidence of effective generalization without overfitting. The latter can be attributed to the MPFA-optimal regularization hyperparameter (e.g., dropout rate) and the inherent architectural resilience of Inception-V4, which incorporates methods such as batch normalization and residual connections. This rapid reduction in loss in the early epochs reflects the efficient gradient flow facilitated by these residual connections, and the subsequent plateau indicates that the model has captured the most salient features in the dataset and reached a stable intersection point.

The accuracy-computational overhead trade-off can be understood as the analysis conducted in Figure 12 positions the MPFA-optimized model well in the design space. A value of high detectability and moderate computational cost is represented by staying in the upper-left corner. The 12 percent higher training time compared to the PFO-optimized one is compensated for by the increased fidelity to detection (+1.82 percent accuracy), which is vital in the security domain, where a single missed intrusion can be disastrous. This trade-off is achieved with the help of the faster and more precise search offered to MPFA, which minimizes the number of fitness evaluations required to find a high-performance model setup.

The superiority of the suggested MPFA-InceptionV4 structure is synthesized in a more comprehensive model comparison and visualized in the radar chart provided in Figure 13. It has an extensive polygon of all five metrics (Accuracy, Precision, Recall, F1-Score, and 100-FAR), indicating a balanced excellence. This is unlike other models, including the 1D-CNN, where a small polygon depicts low performance in multiple dimensions. This is a direct consequence of the multi-objective fitness function (Equation (19)), which steered the MPFA optimization process to achieve the best detector performance in terms of detection rates concurrently, the lowest false alarms, and the management of model complexity, creating a well-rounded detector suitable for use in real-world settings.

Lastly, the analysis of computation time in Table 5 provides a feasible justification for the proposed approach. Although MPFA is slower than an untuned Inception-V4 baseline due to its overhead, its overall runtime is reduced compared to various other meta-heuristic optimizers (WOA, SSA, TLBO, GSA), and its overall accuracy is optimal. This performance is the result of the rapid convergence of MPFA, as shown in Figure 8, which minimizes the number of computationally intensive training cycles required by the neural network in the optimization loop. Consequently, the extra time expenditure over the baseline is tactically justified, as it automates the hyperparameter tuning procedure, substituting months of trial-and-error searches and producing a much more reliable and accurate intrusion detection model, which improves the security stance of smart grid infrastructure.

## 6. Conclusions

This study presents an Inception-V4 model based on the MPFA optimization algorithm as a promising approach for enhancing intrusion detection in intelligent power networks, as demonstrated by its high performance across various evaluation parameters. The combination of the Modified Polar Fox Optimization Algorithm effectively balanced exploration and exploitation, allowing for optimal hyperparameter tuning and leading to a highly accurate and robust detection model. The empirical findings from the smart grid monitoring power dataset validated the model's ability to predict accurately, with F1-scores and accuracy scores reaching the 99.5th percentile mark, in response to various cyber-physical threats, including False Data Injection, Denial-of-Service, and Load Redistribution attacks. This ablation experiment also confirms the synergistic role of the critical cues of MPFA, including learning of courtship and scaling to fitness, in the overall system performance. Furthermore, the framework was characterized by an optimal balance between computational efficiency and detection effectiveness, achieving state-of-the-art results compared to available models and optimizers. The results highlight the promise of metaheuristic-induced deep learning systems in protecting critical cyber-physical infrastructure from changing threats.

**Author Contributions:** Conceptualization, C.T. and L.Z.; methodology, C.T.; software, L.Z.; validation, C.T., L.Z. and H.L.; formal analysis, C.T.; investigation, L.Z.; resources, H.L.; data curation, L.Z.; writing—original draft preparation, C.T.; writing—review and editing, H.L. and C.T.; visualization, L.Z.; supervision, H.L.; project administration, C.T.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by State Grid Sichuan Electric Power Company Science and Technology Program by grant number 52199725001F, with title “Research on Key Technologies of Active Defense and Data Integrity Protection for Power Monitoring Services in New Power Systems”.

**Data Availability Statement:** The data is available in the Power System and Smart Grid Monitoring Power dataset, accessible from the following link: <https://www.kaggle.com/datasets/bachirbarika/power-system> (accessed on 25 December 2025).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Haque, K.A.; Massaoudi, L.; Davis, K.R.; Kabalan, M.; Salmay, H. Cyber-Physical Emulation and Threat Scenario Simulation for Enhanced Microgrid Resilience. *IEEE Access* **2025**, *13*, 101455–101471. [\[CrossRef\]](#)
2. Ootom, S. Risk auditing for Digital Twins in cyber physical systems: A systematic review. *J. Cyber Secur. Risk Audit*. **2025**, *2025*, 22–35. [\[CrossRef\]](#)
3. Tochukwu, I.C.; Nonyelum, O.F.; Misra, S.; Chockalingam, S. Securing mobile edge computing: A survey on cyber-physical threat mitigation for digital sovereignty. *Procedia Comput. Sci.* **2025**, *254*, 211–220. [\[CrossRef\]](#)
4. Bhardwaj, A.; Bharany, S.; Rehman, A.U.; Tejani, G.G.; Hussien, S. Securing cyber-physical robotic systems for enhanced data security and real-time threat mitigation. *EURASIP J. Inf. Secur.* **2025**, *2025*, 1.
5. Barreto, N.E.M.; Aoki, A.R. Cyber-Physical Power System Digital Twins—A Study on the State of the Art. *Energies* **2025**, *18*, 5960. [\[CrossRef\]](#)
6. Varshini, G.S.; Latha, S.; Vikhram, G.R. Impact and Detection of Cyber Attacks in Wide Area Control Application of Cyber-Physical Power System (CPPS). *Comput. Secur.* **2025**, *157*, 104547. [\[CrossRef\]](#)
7. Liu, W. Review of False Data Injection Attacks in Power CPS: Challenges, Detection, and Resilience Strategies. *Preprints* **2025**, *in press*.
8. Zhang, H.; Wang, Z.; Zhang, J. Secure Load Frequency Control of Cyber-Physical Power Systems Under Cyber Attacks and Delays. *IEEE Internet Things J.* **2025**, *12*, 40362–40369.
9. Rouhani, S.H.; Su, C.-L.; Mobayen, S.; Razmjoo, N.; Elsis, M. Cyber resilience in renewable microgrids: A review of standards, challenges, and solutions. *Energy* **2024**, *309*, 133081. [\[CrossRef\]](#)
10. Metwally, A.; Elhenawy, I. Sustainable intrusion detection in vehicular controller area networks using machine intelligence paradigm. *Sustain. Mach. Intell. J.* **2023**, *4*, 1–12. [\[CrossRef\]](#)
11. Walli, S.; Sallam, K. Machine learning for intrusion detection: A reproducible baseline is all you need. *Sustain. Mach. Intell. J.* **2024**, *7*, 1–3. [\[CrossRef\]](#)
12. Tang, Y.; Mishra, S.; Alduaiji, N.; Shukla, P.K.; Yahya, M.; Pang, T. An advanced data analytics approach to a cognitive cyber-physical system for the identification and mitigation of cyber threats in the medical internet of things (MIoT). *J. Supercomput.* **2025**, *81*, 623. [\[CrossRef\]](#)
13. Zografopoulos, I.; Srivastava, A.; Konstantinou, C.; Zhao, J.; Jahromi, A.A.; Chawla, A. Cyber-physical interdependence for power system operation and control. *IEEE Trans. Smart Grid* **2025**, *13*, 2554–2574. [\[CrossRef\]](#)
14. Qudus, L. Resilient systems: Building secure cyber-physical infrastructure for critical industries against emerging threats. *Int. J. Res. Publ. Rev.* **2025**, *6*, 3330–3346. [\[CrossRef\]](#)
15. Kumar, P.; Verma, S. Improving the Resilience of Cyber Security Data Centers to Cyber-Physical Attacks. *Int. J. Manag. Res. Rev.* **2025**, *15*, 14–22.
16. Dayarathne, M.; Jayathilaka, M.; Bandara, R.; Logeeshan, V.; Kumarawadu, S.; Wanigasekara, C. Mitigating Cyber Risks in Smart Cyber-Physical Power Systems Through Deep Learning and Hybrid Security Models. *IEEE Access* **2025**, *13*, 37474–37492. [\[CrossRef\]](#)
17. Kabir, S.; Hannan, N.; Shufian, A.; Zishan, M.S.R. Proactive detection of cyber-physical grid attacks: A pre-attack phase identification and analysis using anomaly-based machine learning models. *Array* **2025**, *27*, 100441. [\[CrossRef\]](#)
18. Ali, O.; Mohammed, O.A. A Review of Multi-Microgrids Operation and Control from a Cyber-Physical Systems Perspective. *Computers* **2025**, *14*, 409.
19. Sadi, M.A.H.; Zhao, D.; Hong, T.; Ali, M.H. Time sequence machine learning-based data intrusion detection for smart voltage source converter-enabled power grid. *IEEE Syst. J.* **2022**, *17*, 2477–2488.
20. Sahani, N.; Zhu, R.; Cho, J.-H.; Liu, C.-C. Machine learning-based intrusion detection for smart grid computing: A survey. *ACM Trans. Cyber-Phys. Syst.* **2023**, *7*, 1–31.
21. Aljohani, A.; AlMuhaini, M.; Poor, H.V.; Binqadhi, H.M. A deep learning-based cyber intrusion detection and mitigation system for smart grids. *IEEE Trans. Artif. Intell.* **2024**, *5*, 3902–3914. [\[CrossRef\]](#)

22. Ankitdeshpandey; Karthi, R. Development of intrusion detection system using deep learning for classifying attacks in power systems. In *Soft Computing: Theories and Applications, Proceedings of the SoCTA 2019, Bengaluru, India, 27–29 December 2019*; Springer: Singapore, 2020; pp. 755–766.
23. Li, X.J.; Ma, M.; Sun, Y. An adaptive deep learning neural network model to enhance machine-learning-based classifiers for intrusion detection in smart grids. *Algorithms* **2023**, *16*, 288.
24. Cavus, M.; Ayan, H.; Sari, M.; Akbulut, O.; Dissanayake, D.; Bell, M. Enhancing Smart Grid Reliability Through Data-Driven Optimisation and Cyber-Resilient EV Integration. *Energies* **2025**, *18*, 4510.
25. Ghiaskar, A.; Amiri, A.; Mirjalili, S. Polar fox optimization algorithm: A novel meta-heuristic algorithm. *Neural Comput. Appl.* **2024**, *36*, 20983–21022. [[CrossRef](#)]
26. Sait, S.M.; Mehta, P.; Yıldız, B.S.; Yıldız, A.R. Artificial neural network-infused polar fox algorithm for optimal design of vehicle suspension components. *Mater. Test.* **2025**, *67*, 1400–1408. [[CrossRef](#)]
27. Yan, X.; Yang, J.; Salami, T. Classification of Indian classical dances using MnasNet architecture with advanced polar fox optimization for hyperparameter optimization. *Sci. Rep.* **2025**, *15*, 18624. [[CrossRef](#)]
28. Nadimi-Shahraki, M.H.; Zamani, H.; Varzaneh, Z.A.; Mirjalili, S. A systematic review of the whale optimization algorithm: Theoretical foundation, improvements, and hybridizations. *Arch. Comput. Methods Eng.* **2023**, *30*, 4113–4159. [[CrossRef](#)]
29. Zhang, H.; Liu, T.; Ye, X.; Heidari, A.A.; Liang, G.; Chen, H.; Pan, Z. Differential evolution-assisted salp swarm algorithm with chaotic structure for real-world problems. *Eng. Comput.* **2023**, *39*, 1735–1769. [[CrossRef](#)] [[PubMed](#)]
30. Díaz, K.Y.G.; De León Aldaco, S.E.; Alquicira, J.A.; Ponce-Silva, M.; Peregrino, V.H.O. Teaching–learning-based optimization algorithm applied in electronic engineering: A survey. *Electronics* **2022**, *11*, 3451.
31. Hashemi, A.; Dowlatshahi, M.B.; Nezamabadi-Pour, H. Gravitational Search Algorithm: Theory; Literature Review and Applications. In *Handbook of AI-Based Metaheuristics*; CRC Press: Boca Raton, FL, USA, 2021; pp. 119–150.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.