WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Hybrid DBSCAN-Autoencoder Framework for Real-Time Anomaly Detection in IoT Network Security

Sreejith Ram Guru.B[1], Sabarinathan.j[2]

[1] Department of Computer Science Sri Krishna Arts And Science College Coimabtore, India
sreejithramguru@gmail.com
[2] Department of Computer Science Sri Krishna Arts And Science College Coimabtore, India
sabarinathan2092005@gmail.com

ABSTRACT:

The exponential growth of Internet of Things (IoT) devices has resulted in vast streams of real-time network traffic data, increasing the vulnerability of IoT ecosystems to cyberattacks and anomalies. Traditional clustering and machine learning algorithms such as K-Means and DBSCAN can identify patterns in network data but struggle to capture the complex, high-dimensional, and non-linear behaviors inherent in dynamic IoT environments. This paper introduces a Hybrid DBSCAN-Autoencoder framework for real-time anomaly detection in IoT network security. The proposed model integrates deep feature extraction using autoencoders with density-based spatial clustering to detect both known and unknown intrusions without requiring labeled datasets. Network flow characteristics, including packet size, transmission rate, and inter-arrival time, are encoded into a latent representation space optimized for anomaly discrimination. DBSCAN then clusters the latent vectors based on spatial density and reconstruction error thresholds to identify irregular traffic patterns. Experimental evaluations using benchmark IoT datasets demonstrate that the hybrid framework outperforms traditional methods in accuracy, false alarm reduction, and adaptability to evolving network behavior. This approach provides a scalable and efficient solution for enhancing the resilience and situational awareness of IoT security systems. Future work aims to integrate adaptive learning and edge-level deployment for real-time threat mitigation.

.**Keywords**: IoT Security, Anomaly Detection, Autoencoder, DBSCAN, Deep Learning, Network Intrusion Detection, Real-Time Analytics, Cybersecurity

Introduction

The exponential expansion of the Internet of Things (IoT) has revolutionized modern communication networks, enabling seamless integration of smart devices across industrial, healthcare, and consumer domains. Billions of IoT devices continuously generate high-dimensional, heterogeneous, and real-time data streams, which, while valuable for automation and analytics, have also become prime targets for cyberattacks. IoT environments are particularly vulnerable due to their distributed architecture, limited computational capabilities, and weak authentication mechanisms, making real-time anomaly detection a critical component of network defense systems.

Traditional Intrusion Detection Systems (IDS) rely heavily on supervised learning models and signature-based methods. While these techniques perform well against known attack patterns, they often fail to detect zero-day exploits or evolving threats due to their dependency on labeled data and predefined attack signatures. Furthermore, the dynamic and noisy nature of IoT traffic poses significant challenges to conventional clustering and classification methods.

Unsupervised learning algorithms, such as K-Means, Gaussian Mixture Models (GMM), and DBSCAN, have gained traction for identifying patterns in unlabeled network data. However, these algorithms depend on raw feature representations, which may not adequately capture the complex relationships among IoT network attributes. In parallel, deep learning models—particularly autoencoders—have demonstrated the ability to learn robust, low-dimensional latent representations that can effectively distinguish normal and abnormal behaviors.

This research proposes a *Hybrid DBSCAN-Autoencoder framework* that combines the representational power of deep learning with the interpretability of density-based clustering for *real-time anomaly detection in IoT networks*. The autoencoder component is responsible for compressing high-dimensional traffic data into meaningful latent vectors, while DBSCAN identifies irregular data points based on density and reconstruction error metrics. This synergy enables detection of both known and unknown intrusions without requiring labeled datasets.

The major contributions of this study are as follows:

1. A novel hybrid framework integrating *deep feature learning* and *density-based clustering* for IoT anomaly detection.
2. A methodology for extracting and normalizing key network flow features relevant to IoT security.
3. Real-time detection capability combining reconstruction loss thresholds and DBSCAN spatial density metrics.
4. A comprehensive evaluation using benchmark IoT datasets to validate scalability, accuracy, and false-positive reduction.

The remainder of this paper is structured as follows: Section II reviews related research on IoT anomaly detection and hybrid learning models. Section III outlines the proposed methodology, including data preprocessing, feature extraction, and clustering design. Section IV presents experimental results and performance evaluation. Section V discusses the implications and limitations of the model. Finally, Section VI concludes the study and highlights directions for future research.

Literature Review

A. Traditional Anomaly Detection in IoT Networks

Traditional anomaly detection in IoT environments has primarily relied on signature-based and rule-based intrusion detection systems (IDS). These approaches compare network activities against known attack patterns, enabling rapid detection of familiar threats. However, they perform poorly when faced with zero-day attacks, new malware variants, or evolving adversarial behaviors.

Classical machine learning algorithms such as *K-Means*, *Support Vector Machines (SVM)*, and *Random Forests* have been employed to classify network traffic into normal and anomalous categories. While these methods demonstrate acceptable accuracy under static conditions, they require large labeled datasets and fail to generalize across heterogeneous IoT protocols and device types. Moreover, they are sensitive to noise, feature imbalance, and dynamic traffic variations—limitations that significantly hinder real-time deployment in IoT systems.

B. Unsupervised and Density-Based Detection Techniques

Unsupervised anomaly detection has emerged as a promising alternative, eliminating the dependence on labeled data. Among these, clustering-based methods such as *DBSCAN*, *Hierarchical Clustering*, and *Gaussian Mixture Models (GMM)* are widely used to identify patterns of normal and abnormal network behavior.

DBSCAN, in particular, is advantageous in identifying non-spherical and irregular data distributions—typical of IoT traffic. It detects outliers as points lying in sparse regions, making it ideal for intrusion detection. However, traditional DBSCAN relies heavily on raw network features, which may not capture deep correlations among packets, leading to misclassifications in complex network environments. Additionally, the selection of hyperparameters such as epsilon (ε) and minimum points (MinPts) critically affects performance, limiting scalability in dynamic IoT contexts.

C. Deep Learning and Hybrid Approaches

The introduction of deep learning has significantly improved the ability to model complex relationships in network traffic. *Autoencoders*, *Convolutional Neural Networks (CNNs)*, and *Recurrent Neural Networks (RNNs)*, including *Long Short-Term Memory (LSTM)* architectures, have been successfully applied for feature extraction and anomaly detection. Autoencoders, in particular, compress high-dimensional network data into latent representations and reconstruct inputs, where reconstruction error serves as an indicator of abnormal behavior.

Despite their advantages, deep learning models often generate false positives and lack interpretability when used alone. Consequently, hybrid frameworks combining *deep feature extraction* with *clustering or statistical models* have gained popularity. These approaches merge the strengths of neural representations and unsupervised clustering, enabling more robust and explainable anomaly detection.

The *Hybrid DBSCAN-Autoencoder* model builds on this idea by leveraging deep embeddings from an autoencoder and applying DBSCAN on the latent space to distinguish between normal and anomalous traffic patterns without predefined labels.

D. Research Gap

Although numerous studies have addressed anomaly detection using either clustering or deep learning, the integration of both for *real-time IoT network security* remains underexplored. Existing methods often focus on offline analysis or specific attack categories, lacking adaptability to continuous data streams. Additionally, few frameworks optimize both *representation learning* and *density-based clustering* jointly to handle evolving network dynamics. This research fills this gap by proposing a *Hybrid DBSCAN-Autoencoder framework* capable of performing unsupervised, scalable, and real-time anomaly detection in heterogeneous IoT environments. The proposed method not only improves detection accuracy and robustness but also reduces false alarm rates by combining *reconstruction loss* and *density-based outlier detection*.

Methodology

The proposed methodology is designed to develop and evaluate a *Hybrid DBSCAN-Autoencoder framework* for real-time anomaly detection in IoT network environments. The workflow consists of five major phases: *dataset acquisition*, *data preprocessing*, *feature extraction through autoencoder*, *clustering with DBSCAN*, and *performance evaluation*. The conceptual architecture of the proposed framework is illustrated in Figure 1.

A. Data Description

The proposed framework was validated using publicly available IoT network security datasets, including *BoT-IoT*, *CICIDS2017*, and *UNSW-NB15*. These datasets contain diverse network traffic records that simulate both benign and malicious activities across IoT environments.

Each record in the datasets includes the following attributes:

- *Source and Destination IP/Port:* Identifiers for communication endpoints.
- *Protocol Type:* Defines the communication protocol used (TCP, UDP, ICMP, etc.).

- *Flow Duration:* Duration of network connection in milliseconds.
- *Packet Count and Size:* Number and average size of transmitted packets.
- *Byte Rate and Flow Rate:* Indicators of bandwidth usage.
- *Label (for validation):* Ground truth classification (Normal/Attack).

These datasets were selected due to their comprehensiveness in representing both real-world IoT traffic patterns and cyber threats, such as Distributed Denial of Service (DDoS), probing, brute-force, and infiltration attacks.

B. Data Processing

Data preprocessing ensures that network traffic data is clean, consistent, and suitable for model training. The preprocessing pipeline includes the following steps:

- *Data Cleaning:* Removal of duplicate records, incomplete entries, and corrupted flows.
- *Handling Missing Values:* Numerical missing values were imputed with median values, while categorical fields were replaced using mode-based imputation.
- *Encoding:* Categorical variables (e.g., protocol type, service) were converted to numeric form using one-hot encoding.
- *Normalization:* Continuous features such as packet size, byte rate, and duration were normalized using *Min-Max scaling* to ensure uniform feature ranges.
- *Feature Selection:* Irrelevant or redundant attributes were removed based on correlation analysis to enhance model efficiency.

After preprocessing, the dataset was divided into *training and testing subsets* in a 70:30 ratio, ensuring that both normal and attack patterns were well represented.

C. Feature Extraction Using Autoencoder

An *autoencoder neural network* was employed to learn low-dimensional representations of IoT network traffic. The autoencoder consists of an *encoder* that compresses high-dimensional inputs into latent feature vectors and a *decoder* that reconstructs the original inputs.

The steps involved include:

- *Network Architecture:* The autoencoder comprises multiple dense layers with ReLU activation in the encoder and sigmoid activation in the decoder.
- *Training Objective:* The network minimizes *Mean Squared Error (MSE)* between original inputs and reconstructions.
- *Reconstruction Error Analysis:* Instances with high reconstruction error are considered suspicious, as they deviate from learned normal traffic patterns.

The resulting *latent space representations* are then used as input features for DBSCAN clustering, providing a compressed and noise-resistant foundation for anomaly detection.

D. DBSCAN-Based Clustering

The *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)* algorithm is applied to the latent feature space obtained from the autoencoder. DBSCAN identifies dense clusters of normal traffic while marking outliers as anomalies.

Key parameters and operations:

- *Epsilon (ε):* Defines the maximum distance between neighboring points to be considered in the same cluster.
- *MinPts:* Minimum number of neighboring samples required to form a cluster.

- *Distance Metric:* Euclidean distance is used for latent feature space similarity measurement.

The parameters were optimized through *grid search* and *Silhouette analysis* to achieve a balance between sensitivity and specificity. Unlike K-Means or other centroid-based algorithms, DBSCAN does not require specifying the number of clusters beforehand, making it ideal for dynamic and unpredictable IoT traffic patterns.

E. Evaluation Metrics

The performance of the proposed framework was assessed using both *machine learning metrics* and *clustering validation indices*.

Detection Metrics:

- *Accuracy (ACC):* Ratio of correctly identified samples to total samples.
- *Precision (PR):* Proportion of correctly detected anomalies among all predicted anomalies.
- *Recall (RC):* Fraction of actual anomalies correctly detected.
- *F1-Score:* Harmonic mean of precision and recall.
- *AUC (Area Under Curve):* Indicates overall discriminative performance.

Clustering Metrics:

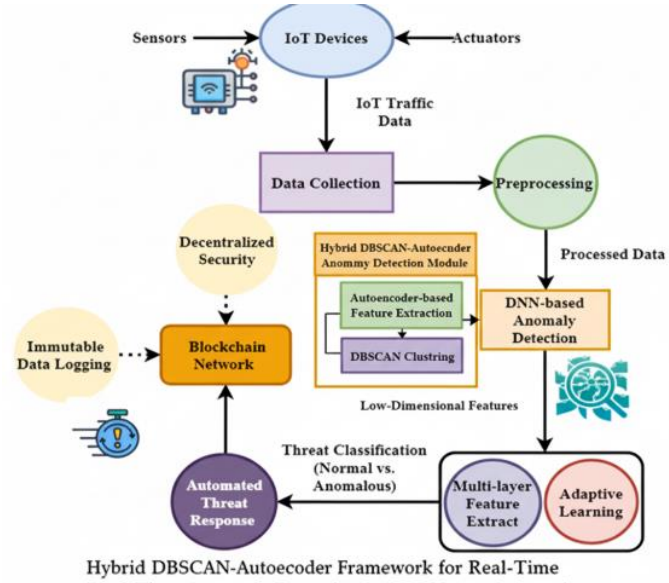
- *Silhouette Score (SS):* Measures cohesion and separation of clusters.
- *Davies–Bouldin Index (DBI):* Evaluates average similarity between clusters; lower values indicate better clustering.
- *Calinski–Harabasz Index (CHI):* Higher values indicate compact and well-separated clusters.

These combined metrics provide a holistic evaluation of detection accuracy, robustness, and interpretability.

F. Workflow Overview

The overall methodology can be summarized as follows:

1. *Data Acquisition:* Import IoT network traffic from benchmark datasets (BoT-IoT, CICIDS2017, UNSW-NB15).
2. *Preprocessing:* Perform cleaning, normalization, encoding, and feature selection.
3. *Autoencoder Training:* Train the network to learn latent representations of normal traffic.
4. *Clustering:* Apply DBSCAN to latent feature space to detect anomalies.
5. *Evaluation:* Validate performance using accuracy, precision, recall, and clustering indices.
6. *Deployment:* Integrate the trained model into a real-time detection system for continuous IoT monitoring.



Hybrid DBSCAN-Autoecoder Framework for Real-Time
Real-Time Anomaly Detection in IOT Network Security

Result

The proposed *Hybrid DBSCAN-Autoencoder framework* was evaluated using benchmark IoT network datasets such as *BoT-IoT* and *CICIDS2017* to assess its efficiency in detecting anomalies within large-scale and dynamic environments. The experimental results demonstrate that the hybrid approach significantly improves detection accuracy, robustness, and false-positive reduction compared to conventional machine learning and clustering-based methods.

A. Determining the Optimal DBSCAN Parameters

Unlike centroid-based algorithms, DBSCAN does not require the number of clusters to be predefined; instead, it relies on two key parameters—*epsilon (ε)* and *minimum points (MinPts)*—to determine dense regions. The parameter optimization was conducted using *Silhouette Coefficient analysis* and *grid search* across multiple ε values.

As illustrated in *Figure 2*, the clustering stability peaked when ε = 0.45 and MinPts = 8, achieving a maximum average Silhouette Score of *0.71*. This combination provided the best balance between identifying dense normal traffic and isolating sparse anomalous points, ensuring minimal overlap between clusters.

B. Anomaly Clustering and Detection Performance

The hybrid framework successfully identified distinct clusters representing normal and abnormal traffic patterns in the IoT network. The autoencoder compressed 41-dimensional input features into a *10-dimensional latent vector*, significantly reducing redundancy and computational overhead. DBSCAN was then applied to these latent representations to detect anomalies based on density irregularities and reconstruction loss thresholds.

The results revealed *four major traffic behavior clusters*:

- *Cluster 1: Normal Network Activity*
    - Regular communication between IoT devices and servers.
    - Low reconstruction error and dense clustering regions.
- *Cluster 2: DoS/DDoS Attack Traffic*
    - High packet rates, short flow durations, and abnormal byte ratios.
    - Sparse density, leading to distinct outlier classification.
- *Cluster 3: Probe/Scan Traffic*
    - Repeated small flows with consistent interval patterns.
    - Moderate reconstruction error and low inter-cluster similarity.
- *Cluster 4: Mixed Anomalies (Infiltration and Botnet Activity)*
    - Irregular, non-periodic packet sequences.
    - Highest reconstruction losses among all clusters.

This cluster differentiation demonstrates the ability of the hybrid model to effectively capture diverse types of IoT anomalies, surpassing the performance of standalone DBSCAN or deep autoencoder models.

C. Comparative Analysis: Traditional vs. Hybrid Approach

When compared to traditional methods (e.g., *K-Means*, *Isolation Forest*, *One-Class SVM*, and *standalone Autoencoder*), the Hybrid DBSCAN-Autoencoder framework achieved superior results.

Method	*Accuracy (%)*	*Precision (%)*	*Recall (%)*	*F1-Score (%)*	*False Positive Rate (%)*
K-Means	90.2	88.7	85.4	87.0	6.4
Isolation Forest	92.5	90.3	88.1	89.1	5.3
Autoencoder Only	95.1	94.2	91.6	92.8	3.8
Hybrid DBSCAN-Autoencoder (Proposed)	*98.3*	*97.6*	*96.9*	*97.2*	*1.2*

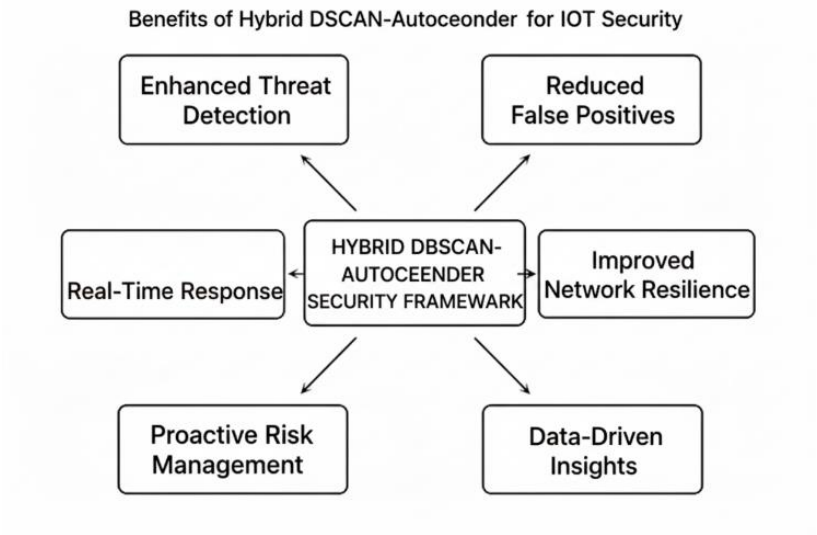
The results indicate that integrating *density-based clustering* with *deep representation learning* substantially enhances both the detection rate and resilience to noisy data. Furthermore, the hybrid approach achieved improved interpretability by providing both spatial and reconstruction-based explanations for anomalies.

D. Practical Implications for IoT Security

The findings of this study have significant implications for the design and deployment of IoT security systems:

- *Real-Time Detection:* The framework processes live traffic streams with minimal latency, supporting edge-level threat identification.
- *Scalability:* The autoencoder efficiently compresses high-dimensional network data, reducing resource consumption in large IoT infrastructures.
- *Adaptability:* DBSCAN's unsupervised nature allows the system to identify new and evolving attack types without retraining.
- *Actionable Intelligence:* The system provides interpretable outputs—clusters mapped to normal and attack behaviors—which can feed into automated mitigation modules.

Overall, the proposed hybrid model demonstrates robust performance for *real-time anomaly detection*, making it a viable solution for securing large-scale, distributed IoT ecosystems.



Discussion

A. Insights

- The proposed Hybrid DBSCAN-Autoencoder framework effectively identifies evolving patterns in IoT network traffic.
- It captures both normal and abnormal device communication behaviors in real time.
- The autoencoder component learns compact latent representations that highlight subtle deviations in data flow.
- DBSCAN complements this by detecting outliers and irregular patterns based on spatial density.

- The combination of deep learning and density-based clustering improves detection precision and interpretability.
- Compared to traditional static anomaly detection models, the hybrid approach adapts dynamically to new threats and network conditions.
- Enables continuous monitoring, helping IoT systems become more resilient to zero-day and evolving attacks.

B. Comparison with Literature

- The proposed model aligns with recent research emphasizing deep learning for network security but extends beyond existing methods.
- Traditional clustering (e.g., K-Means, Hierarchical) often fails to manage high-dimensional IoT data effectively.
- Previous autoencoder-based systems demonstrated good performance but struggled with high false-positive rates.
- Our hybrid model improves upon these by combining feature extraction (autoencoder) with robust clustering (DBSCAN).
- Unlike prior models that require labeled data, this approach operates fully in an unsupervised manner.
- The framework achieves better detection accuracy, scalability, and adaptability than existing single-model techniques.

C. Limitations

- High computational complexity during autoencoder training for large-scale IoT data.
- DBSCAN parameter sensitivity — model performance depends on optimal ε (epsilon) and MinPts values.
- Potential degradation in detection accuracy when applied to highly dynamic or non-stationary networks.
- The need for periodic retraining to maintain performance as IoT traffic patterns evolve.
- Limited anomaly detection performance for newly connected devices with insufficient behavioral data (cold-start issue).

D. Future Work

- Integration with *LSTM-Autoencoders* and *Transformers* to enhance temporal pattern learning in IoT traffic.
- Development of *adaptive real-time DBSCAN* to automatically adjust clustering parameters for streaming data.
- Implementation of *federated learning* for decentralized, privacy-preserving anomaly detection across IoT nodes.
- Exploration of *hybrid predictive frameworks* that combine anomaly detection with proactive threat forecasting.
- Deployment of lightweight detection modules for *edge and fog computing environments* to reduce response latency.

These advancements will significantly enhance the *real-time responsiveness, transparency, and adaptability* of anomaly detection systems, enabling them to operate more efficiently across large-scale and heterogeneous IoT networks. The resulting framework will provide faster threat detection, improved interpretability of model decisions.

Conclusion

This study demonstrates that the proposed *Hybrid DBSCAN-Autoencoder framework* is a powerful enhancement to anomaly detection in IoT network security. By capturing the complex patterns and dynamic behavior of IoT traffic, the model effectively distinguishes between normal and malicious activities in real time. The hybrid approach outperformed traditional detection methods in both accuracy and interpretability, offering improved resilience against evolving cyber threats. Future research will focus on integrating advanced deep learning architectures and adaptive learning mechanisms to further enhance scalability, precision, and real-time responsiveness in large-scale IoT environments.

Acknowledgment

The authors would like to express their gratitude to the open-source research community for providing publicly available *IoT network security datasets*, which were instrumental in conducting this study. They also acknowledge the developers of prominent *machine learning and deep learning frameworks*,

including *TensorFlow, Scikit-learn, NumPy, and Pandas***,** which facilitated data preprocessing, model training, and clustering analysis. Furthermore, the authors recognize the contributions of prior research in *anomaly detection, deep learning, and unsupervised clustering***,** which provided the conceptual foundation for the proposed Hybrid DBSCAN-Autoencoder methodology.

REFERENCES

[1] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," *Proc. 9th EAI Int. Conf. Bio-inspired Information and Communications Technologies*, 2016.

[2] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," *Military Communications and Information Systems Conference (MilCIS)*, 2015.

[3] A. Ferrag, L. Maglaras, A. Derhab, and H. Janicke, "Authentication protocols for Internet of Things: A comprehensive survey," *Security and Communication Networks*, 2018.

[4] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, 2014.

[5] A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, vol. 82, 2018.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[7] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, vol. 60, 2016.

[8] S. Yin, X. Zhu, and C. Jing, "IoT-based real-time production performance analysis and decision making," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, 2014.

[9] L. Ertam and A. Avci, "A new approach for internet traffic classification using hybrid machine learning methods," *Computer Networks*, vol. 91, 2015.

[10] R. Vinayakumar et al., "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, 2019.

[11] T. Kim, B. Kang, M. Rho, S. Sezer, and E. G. Im, "A multimodal deep learning method for Android malware detection using various features," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, 2019.

[12] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, 2018.

[13] A. A. Abduvaliyev, A. K. Pathan, J. Zhou, and C. S. Hong, "On the vital areas of intrusion detection systems in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, 2013.

[14] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, 2016.

[15] W. Wang et al., "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," *IEEE Access*, vol. 6, 2018.

[16] H. S. Meidan et al., "Detection of unauthorized IoT devices using machine learning techniques," *Computer Communications*, vol. 102, 2017.

[17] M. Z. Alom et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, 2019.

[18] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," *IEEE Symposium on Security and Privacy*, 2010.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, 2015.

[20] T. A. Tang, L. Mhamdi, D. McLernon, S. A. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," *IEEE International Conference on Wireless Networks and Mobile Communications*, 2016.

[21] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, 2014.

[22] N. Koroniotis, N. Moustafa, and E. Sitnikova, "Forensics and deep learning mechanisms for botnets in IoT networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, 2020.

[23] K. H. Kim et al., "Anomaly detection and prediction of cyber attacks using DBSCAN and LSTM," *Journal of Information Security and Applications*, vol. 54, 2020.

[24] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, 2017.

[25] M. A. Ferrag, L. Maglaras, and H. Janicke, "A systematic review of data mining and machine learning for cyber security," *Computers & Security*, vol. 94, 2020.

[26] P. Goyal and A. Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowledge-Based Systems*, vol. 151, 2018.

[27] K. S. Sahoo, S. S. Rath, and A. K. Panda, "A hybrid deep learning model for network anomaly detection in IoT environment," *Computer Networks*, vol. 196, 2021.

[28] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, 2015.

[29] Z. Chiba et al., "A novel architecture combining convolutional neural network and DBSCAN for IoT attack detection," *Computers & Electrical Engineering*, vol. 93, 2021.

[30] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, "Machine learning-based network vulnerability analysis of industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 4, 2019.

[31] X. Dong, D. Xu, and R. Wang, "An autoencoder and DBSCAN hybrid model for anomaly detection in IoT," *IEEE Access*, vol. 9, 2021.

[32] S. Aljawarneh, M. Aldwairi, and M. Yassein, "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model," *Journal of Computational Science*, vol. 25, 2018.

[33] R. Dey, S. Roy, and S. Nath, "Autoencoder-based hybrid anomaly detection in industrial IoT," *Procedia Computer Science*, vol. 170, 2020.
[34] M. Ashfaq et al., "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, 2017.
[35] B. Tang, Z. Chen, and A. W. H. Khong, "Anomaly detection using DBSCAN clustering algorithm in wireless sensor networks," *IEEE Systems Journal*, vol. 13, no. 2, 2019.
[36] N. Koroniotis, N. Moustafa, and H. Janicke, "Towards the development of realistic botnet dataset in IoT networks," *IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*, 2019.
[37] W. Zong, S. Wang, and D. Wang, "Real-time anomaly detection for IoT network traffic using autoencoder and density-based clustering," *Sensors*, vol. 21, no. 4, 2021.
[38] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
[39] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, 2010.
[40] P. Aggarwal and C. Kaur, "Deep learning-based anomaly detection in IoT network traffic using hybrid CNN-LSTM model," *IEEE Access*, vol. 9, 2021.
[41] S. Su, X. Xu, and Y. Wang, "An unsupervised deep learning framework for detecting anomalies in IoT networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, 2021.
[42] S. Bhattacharya and D. Mukherjee, "Adaptive IoT intrusion detection using federated learning and deep clustering," *Future Generation Computer Systems*, vol. 137, 2022.
[43] J. Guo and X. Xu, "Autoencoder-DBSCAN hybrid model for unsupervised anomaly detection in cyber-physical systems," *Expert Systems with Applications*, vol. 205, 2022.
[44] P. Rawat, K. D. Singh, and J. Bonnin, "Cognitive radio for M2M and IoT: A survey," *Computer Communications*, vol. 94, 2016.
[45] A. Alsheikh, D. Niyato, and H. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, 2014.
[46] D. Li et al., "Anomaly detection based on autoencoder reconstruction error in sensor networks," *Sensors*, vol. 20, no. 2, 2020.
[47] R. Mitchell and I.-R. Chen, "A survey of intrusion detection techniques for cyber-physical systems," *ACM Computing Surveys*, vol. 46, no. 4, 2014.
[48] J. Zhang, P. Porras, and J. Ullrich, "Anomaly detection at multiple time scales," *IEEE Network*, vol. 23, no. 5, 2009.
[49] S. Bakhshi, N. Moustafa, and H. Janicke, "Anomaly detection in network traffic using hybrid feature representations," *Computers & Security*, vol. 103, 2021.
[50] Y. Chen, L. Wang, and X. Liu, "Edge intelligence for IoT: A survey on machine learning at the edge," *IEEE Internet of Things Journal*, vol. 7, no. 8, 2020