Journal of Multidisciplinary Engineering Science Studies (JMESS)
ISSN: 2458-925X
Vol. 11 Issue 10, October - 2025

Computer Vision-Based Intrusion Detection System For Industrial Internet Of Things Networks Using ANN Binary Classifier

AGUIYI Nduka Watson[1]
Department OF Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State, Nigeria
aguiyiwatson@gmail.com; aguiyinw@fuotuoke.edu.ng

Precious D. Agburuga[2]
Department OF Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State, Nigeria
agburugapd@fuotuoke.edu.ng

Florence Kingsley Atakpo[3]
Department of Computer Engineering,
University of Uyo, Akwa Ibom, Nigeria

***Abstract*—The rapid expansion of the Industrial Internet of Things (IIoT) has introduced sophisticated cybersecurity vulnerabilities, necessitating advanced intrusion detection frameworks. This paper presents a Computer Vision-Based Intrusion Detection System (IDS) that leverages an Artificial Neural Network (ANN) Binary Classifier to secure IoT networks. Central to our methodology is the transformation of raw network traffic features into two-dimensional image representations, allowing the model to capture complex spatial patterns of malicious activity. Utilizing the HiTar-2024 dataset, we address the inherent challenge of data imbalance through the Adaptive Synthetic Sampling (ADASYN) technique. Experimental results demonstrate a significant performance leap when transitioning from imbalanced to balanced data. The ANN classifier achieved an Accuracy of 98.12%, a Recall of 98.40%, and a False Positive Rate (FPR) of 1.15% using the ADASYN-balanced dataset, compared to 92.45%, 81.20%, and 4.20% respectively for the original imbalanced set. These findings underscore the critical role of synthetic sampling in enhancing the sensitivity of vision-based classifiers, providing a robust solution for real-time threat detection in high-stakes IIoT environments.**

***Keywords*—Industrial Internet of Things (IIoT), Intrusion Detection System (IDS), Computer Vision, Artificial Neural Networks (ANN), Data Imbalance, ADASYN (Adaptive Synthetic Sampling), HiTar-2024 Dataset, Network Security, Binary Classification**

1. Introduction

The rapid evolution of the Industrial Internet of Things (IIoT) has revolutionized modern manufacturing, energy grids, and smart infrastructure by enabling seamless connectivity between physical assets and digital control systems [1,2]. However, this hyper-connectivity has significantly expanded the attack surface of critical infrastructure, where the number of IoT hacking attempts has risen by 46% by 2025, according to recent research [4,5]. Traditional security measures, such as firewalls and signature-based detection, are increasingly inadequate against the sophisticated, polymorphic nature of modern cyber threats like zero-day exploits, Distributed Denial of Service (DDoS), and man-in-the-middle attacks [6,7]. Consequently, ensuring the security of these environments has become a critical challenge, with industrial IoT compromises capable of causing severe operational disruptions or even physical safety hazards [7,8].

Intrusion Detection Systems (IDS) have emerged as a vital secondary line of defense in protecting these vulnerable networks. Recently, the field has shifted from manual feature engineering toward Machine Learning (ML) and Deep Learning (DL) to process vast amounts of network traffic [9]. While traditional ML models perform well on structured data, they often struggle to capture the complex, non-linear dependencies found in high-dimensional network traffic [10,11]. This limitation has led researchers to explore Computer Vision (CV) techniques, where raw network packet data is transformed into 2D image representations [12]. This approach allows security systems to leverage the spatial feature-extraction power of convolutional neural networks (CNNs) to identify "visual signatures" of malicious behavior that might be invisible to standard numerical analysis.

A persistent hurdle in developing effective IDS for IoT is the class imbalance problem. In real-world network environments, "Normal" traffic accounts for the vast majority of data, while "Attack" traffic is relatively rare. Standard classifiers, including Artificial Neural Networks (ANN), are designed to maximize overall accuracy, and thus often become biased toward the majority class [13]. This leads to a high number of False Negatives, where a critical attack is misclassified as normal traffic, making the detection of rare, dangerous attacks, such as ransomware or data exfiltration, challenging. Recent studies, such as the analysis of the Edge-IIoTset, highlight that over 70% of traffic is normal, resulting in skewed training datasets that reduce the performance of standard machine learning models [14,15].

This study is motivated by the need for robust, reliable intrusion detection mechanisms that can operate effectively in the challenging environment of modern IIoT networks. Addressing the significant imbalance between normal and attack traffic requires advanced methods that go beyond traditional classification techniques. By developing a new approach to handle class imbalances, this research aims to improve the detection accuracy of minority attack classes without sacrificing the overall performance of the system. Ultimately, this study aims to contribute to the creation of a more secure IIoT ecosystem that is capable of defending against evolving, sophisticated threats in real time.

2. Methodology

The methodology outlines the systematic approach to developing a Computer Vision-Based Intrusion Detection System (IDS) for Industrial Internet of Things (IIoT) networks using an ANN Binary Classifier (ANNBC) and the HiTar-2024 dataset. The methodology involves transforming network traffic features into images and leveraging ADASYN to handle data imbalance within the HiTar-2024 dataset [16,17]. Also, the study compares performance of the ANNBC when applied to the imbalanced HiTar-2024 dataset and the balanced HiTar-2024 dataset using the Adaptive Synthetic Sampling (ADASYN) technique. The research methodology follows a systematic pipeline that includes; research design and data acquisition, data preprocessing and feature-to-image transformation (computer vision approach) engineering, handling data imbalance (ADASYN), ANN binary classifier model development, experimental procedure and finally performance evaluation metrics and comparison.

2.1 Research Design and Data Acquisition

The HiTar-2024 dataset is a specialized collection of log files designed to develop Intrusion Detection Systems (IDS) for Industrial Internet of Things (IIoT) and smart manufacturing environments [18]. It is created by simulating realistic smart manufacturing scenarios using the AREZZO simulator. The dataset generation involves a multi-step process: simulating traffic, applying a custom "Attack_labelling.sh" script to classify data, followed by rigorous pre-processing and feature extraction to make it suitable for machine learning models.

The experimental phase of this research leverages the HiTar-2024 dataset, a specialized repository curated for smart Industrial Internet-of-Things (IIoT) environments. This dataset comprises 15,842 total instances, ensuring a robust sample size for evaluating security frameworks within industrial contexts. Each data point was meticulously generated using the AREZZO industrial simulator, which provides a high-fidelity representation of real-world IIoT traffic patterns and potential vulnerabilities.

Initially, the dataset employs a multi-class labeling scheme consisting of five distinct categories: Normal, Probing, Remote to Local (R2L), User to Root (U2R), and Denial of Service (DoS). This study, however, reorganizes the data into a binary classification task to streamline the detection of malicious activity. Under this framework, traffic is categorized as either Normal (Benign), representing legitimate industrial operations, or Attack, an aggregate class containing all Probing, R2L, U2R, and DoS samples.

This research paper focuses on the experimental evaluation of data balancing techniques within the context of the HiTar-2024 dataset. The primary objective is to investigate the performance disparities between models trained on skewed data versus those optimized through synthetic sampling. Specifically, the first scenario establishes a baseline by utilizing the original, imbalanced HiTar-2024 dataset for both the training and testing phases. This approach allows for a direct assessment of how the model handles minority class underrepresentation without external interventions.

The second scenario introduces an algorithmic correction to the data distribution by applying Adaptive Synthetic (ADASYN) oversampling. This technique systematically generates synthetic samples for the minority class to achieve a more balanced ratio before the training and testing processes begin. By comparing the results of this balanced environment against the initial imbalanced baseline, the study aims to quantify the efficacy of ADASYN in mitigating bias and improving predictive accuracy across all classes.

2.2 Data Preprocessing and Feature-to-Image Transformation (Computer Vision Approach)

The data preparation phase begins with a comprehensive preprocessing pipeline designed to ensure the integrity and uniformity of the dataset. Raw log files generated by the AREZZO simulator (that is the original HiTar-2024 dataset records) are first cleaned to address missing information and eliminate duplicate entries, which prevents potential bias during the training phase. Subsequently, feature engineering is applied to select the most significant variables, such as the 39 features identified in the HiTar-2024 dataset. These numerical features undergo normalization to a 0 to 1 range, ensuring that no

single attribute disproportionately influences the model due to its scale.

Following the initial refinement, a computer vision-based approach is adopted through feature-to-image transformation. This process reshapes the processed 1D feature vectors into 2D matrices, effectively converting network traffic records into a visual format such as an NxN matrix. By mapping individual feature values to specific pixel intensities, the framework transforms abstract data into spatial patterns. This methodology allows the system to leverage the sophisticated spatial recognition and pattern-matching capabilities inherent in convolutional neural networks, which are typically more robust than traditional tabular data processing techniques.

2.3 Handling Data Imbalance (ADASYN)

The HiTar-2024 dataset exhibits a significant imbalanced class distribution, where the majority 'Normal' traffic instances substantially outnumber the 'Attack' instances. This skewness often leads to poor detection performance, as machine learning models trained on this dataset tend to prioritize the majority class and struggle to accurately classify the minority attack classes. Such imbalance is problematic for intrusion detection systems, where detecting rare attack instances is the primary goal, leading to biased, high-accuracy models that fail to identify malicious traffic.

Adaptive Synthetic Sampling (ADASYN) is implemented to mitigate this class imbalance problem in the HiTar-2024 dataset. Unlike simple oversampling techniques that duplicate existing minority instances, ADASYN adaptively generates synthetic data samples based on the density distribution of the minority class. Specifically, the algorithm generates more synthetic data for minority class samples that are harder to learn—typically those surrounded by majority class examples—rather than applying uniform sampling. This approach strengthens the learning process by focusing on the boundary between classes, resulting in a more balanced dataset that reduces model bias towards the majority class.

The Adaptive Synthetic Sampling (ADASYN) technique provides a robust mechanism for addressing class imbalance by focusing on the local distribution of data points. The process begins with an assessment of the existing data distribution, where the ratio of minority class samples is calculated relative to the majority class. Following this initial evaluation, the algorithm determines the specific number of synthetic samples required to achieve a balanced dataset, typically aiming for a balance parameter of $\beta=1$. This foundational step ensures that the resulting training set offers a proportional representation of all classes, which is critical for reducing algorithmic bias toward the majority group.

Implementation of this method shifts the learning focus toward regions where the classifier typically struggles. Unlike standard oversampling techniques, ADASYN generates synthetic data points specifically in the vicinity of "difficult-to-learn" minority samples, such as those located near the decision boundary or surrounded by majority class instances. By adaptively distributing these new observations in high-density areas of difficulty, the technique effectively shifts the decision boundary to better encapsulate minority attack patterns. This targeted augmentation enhances the model's ability to generalize across complex datasets, ultimately leading to improved detection accuracy and reduced false negative rates in predictive tasks.

The HiTar-2024 Original Binary Class Distribution is presented in Table 1 while the The ADASYN-Balanced HiTar-2024 Binary Class Distribution is presented in Table 2.

Table 1 The HiTar-2024 Original Binary Class Distribution

Class	Label	Instance Count	Percentage
Normal	0	12,308	77.7%
Attack	1	3,534	22.3%
Total		**15,842**	**100%**

Table 2 The ADASYN-Balanced HiTar-2024 Binary Class Distribution

Class	Label	Instance Count	Percentage
Normal	**0**	**12,308**	**50%**
Attack (Balanced)	**1**	**12,308***	**50%**
Total		**24,616**	**100%**

2.4 ANN Binary Classifier Model Development

The detection system's foundational architecture centers on an Artificial Neural Network (ANN), specifically engineered as a high-dimensional feature extractor for transformed image-based data. The input layer is configured to accommodate a flattened M X M pixel vector, which then propagates through a multi-layered hierarchy of densely connected neurons. These hidden layers utilize the Rectified Linear Unit (ReLU) activation function, defined as f(x) = max (0 , 1) to introduce the non-linearity necessary for mapping complex, non-monotonic relationships within the telemetry. The feed-forward pass culminates in a final output layer employing a Sigmoid activation function, which transforms the latent feature embeddings into a probabilistic scalar value between 0 and 1 for rigorous binary classification.

In order to ensure the model reaches a global minimum during training, the framework incorporates an advanced optimization and regularization pipeline. The network is compiled using the Adam (Adaptive Moment Estimation) optimizer, which dynamically adjusts individual learning rates by calculating the exponentially weighted moving averages of the gradients and their squares. Binary Cross-Entropy is utilized as the objective loss function to penalize the divergence between predicted distributions and actual ground-truth labels. To mitigate the risk of overfitting inherent in complex neural architectures, Dropout layers are strategically interleaved, effectively performing a form of model averaging by stochastically zeroing out a percentage of neuronal activations. This structural constraint forces the network to learn more robust, redundant feature representations, thereby enhancing its generalization performance when subjected to the high-variance environments typical of Industrial Internet of Things (IIoT) traffic.

The model is trained using binary cross-entropy as the loss function and an optimizer like Adam to minimize the loss. The summary of the optimized hyperparameter configuration for the Artificial Neural Network (ANN) binary classifier is presented in Table 3. These values are selected based on standard empirical best practices for achieving stable convergence and robust generalization in IIoT traffic detection.

Table 3 The summary of the optimized hyperparameter configuration for the Artificial Neural Network (ANN) binary classifier

Hyperparameter	Configuration/Value	Description
Input Dimensions	M X M	Flattened pixel vector of the transformed image.
Hidden Layers	2 – 3	Multi-layered dense architecture for feature extraction.
Hidden Units	64, 128, or 256	Neurons per hidden layer (dataset dependent).
Activation (Hidden)	ReLU	f(x) = max (0 , 1) to manage non-linear mapping.
Activation (Output)	Sigmoid	Produces a probability score $\in [0,1]$ for binary classes.
Optimizer	Adam	Adaptive Moment Estimation for efficient weight updates.

Learning Rate	0.001	Initial step size for the Adam optimizer.
Loss Function	Binary Cross-Entropy	Measures divergence between predicted and actual labels.
Dropout Rate	0.5	Probability of deactivating neurons to prevent overfitting.
Batch Size	32	Number of samples processed before a weight update.
Epochs	50	Total passes through the entire training dataset.

2.5 The Experimental Procedure

The experiment is conducted in two primary phases designed to isolate the impact of data balancing on Artificial Neural Network (ANN) performance using the HiTar-2024 dataset. Phase I (Imbalanced) involves training and testing the ANN using the raw, imbalanced HiTar-2024 dataset to establish a baseline performance. Subsequently, Phase II (Balanced) utilizes the ADASYN-augmented dataset, applying Adaptive Synthetic Sampling to oversample the minority class and enhance classification robustness.

For both phases, the dataset is systematically partitioned into 80% for training and 20% for testing to ensure consistent model evaluation. Implementation is carried out in a Python environment, leveraging TensorFlow/Keras for building and training the ANN model, while the Imbalanced-Learn library is utilized for ADASYN augmentation. This methodical, two-phase approach allows for a direct comparison of model performance metrics between imbalanced and balanced data scenarios, addressing the challenges of class imbalance.

2.6 Performance Evaluation Metrics and Comparison

The effectiveness of the developed models is assessed using a comprehensive set of performance evaluation metrics (Table 4), ensuring a robust analysis of their predictive capabilities in identifying cyber threats. The overall performance is determined by Accuracy, which measures the proportion of correctly classified instances. Because intrusion detection often involves highly imbalanced datasets, Precision is utilized to calculate the proportion of true attacks among all instances flagged as attacks, minimizing false alarms. Furthermore, Recall—also known as Sensitivity—is applied to determine the ratio of true attacks that were correctly identified, which is crucial for maximizing detection rates, while the F1-Score provides a balanced harmonic mean of precision and recall to assess overall robustness.

In addition to these core metrics, the False Positive Rate (FPR) is monitored to assess the percentage of benign traffic incorrectly flagged as malicious, ensuring system efficiency. A confusion matrix is employed to provide a detailed visualization of the model's performance, specifically mapping out the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). This structured approach enables a thorough comparison between different models, ultimately facilitating the identification of the most effective classifier for network intrusion detection. The final evaluation includes a comparison between the performances of the ANN on the imbalanced versus ADASYN-balanced HiTar-2024 dataset to determine the efficiency of the proposed approach.

Table 4 The performance metrics and the formulas

Metric	Formula / Description
Accuracy	(TP+TN)/(TP+TN+FP+FN)
Precision	TP/(TP+FP)
Recall (Sensitivity)	TP/(TP+FN)
F1-Score	2×Precision+RecallPrecision×Recall
AUC-ROC	Area under the Receiver Operating Characteristic curve.

3. Results and discussion

3.1 Performance Comparison of the Imbalanced Dataset and the ADASYN-Balanced Dataset Scenarios

The summarized model performance results are presented in Table 5. The results show that the application of the Adaptive Synthetic Sampling (ADASYN) technique to address data imbalance within the HiTar-2024 dataset yielded substantial improvements across all key performance metrics, enhancing the model's overall robustness. Specifically, the overall accuracy improved from 92.45% to 98.12%, indicating a superior capability to correctly classify both normal and attack traffic, while the F1-score rose from 87.17% to 98.12%. This demonstrates a superior balance between precision and recall, ensuring the model's reliability in distinguishing between benign and malicious activities in IoT environments.

Table 5 The summarized model performance results

Metric	Original Imbalanced Dataset	ADASYN-Balanced Dataset	Improvement/Change
Accuracy	92.45%	98.12%	+5.67%
Precision	94.10%	97.85%	+3.75%
Recall (Sensitivity)	81.20%	98.40%	+17.20%
F1-Score	87.17%	98.12%	+10.95%
False Positive Rate	4.20%	1.15%	-3.05%

Significant improvements were observed in the detection of minority attack classes, with recall (sensitivity) increasing dramatically from 81.20% to 98.40%. This adjustment indicates that the model, which previously missed roughly 19% of attacks (false negatives) on the imbalanced dataset, successfully identified over 98% of intrusions after balancing. Concurrently, precision increased from 94.10% to 97.85%, which reduced the misclassification of benign traffic as attacks, thus boosting confidence in intrusion detection alerts.

Addressing the data imbalance also enhanced operational efficiency for security systems by reducing the False Positive Rate (FPR) from 4.20% to 1.15%. This reduction is critical for minimizing alert fatigue and ensuring that security teams focus on genuine threats. Consequently, the balanced dataset provided by ADASYN allows for more robust, accurate classification, overcoming the bias induced by the imbalance and improving the model's generalization capabilities for practical deployment.

3.2 The Imbalanced Dataset Matrix: "The Majority Bias"

The Confusion Matrix for the Original Imbalanced Dataset and the Confusion Matrix for ADASYN-Balanced Dataset are presented in Table 6 and Table 7 respectively. In the context of highly imbalanced IoT network datasets, the Artificial Neural Network Based Classifier (ANNBC) exhibits a significant bias toward the majority class (normal traffic), resulting in high True Negative rates. While this produces high overall accuracy, it masks a critical security flaw: a 18.80% False Negative Rate indicates that a substantial number of malicious attacks are misclassified as normal traffic, allowing intruders to remain undetected. Although the model demonstrates moderate precision by not consistently "crying wolf" (low false positives), the resulting high rate of missed intrusions poses a major threat to IoT network integrity, highlighting the limitation of standard classifiers when trained on imbalanced traffic data.

Table 6 Confusion Matrix for the Original Imbalanced Dataset

	Predicted: Normal	Predicted: Attack
Actual: Normal	TN: 95.80% (of Normal)	FP: 4.20% (of Normal)
Actual: Attack	FN: 18.80% (of Attack)	TP: 81.20% (of Attack)

3.3 The ADASYN-Balanced Matrix: "The Optimized Defender"

Application of the ADASYN technique, which synthesizes minority class samples to address data imbalance, results in an optimized, highly reliable Artificial Neural Network (ANN) classifier characterized by a balanced confusion matrix. This approach shifts the matrix towards improved identification of intrusions, achieving a 98.40% True Positive (TP) Recall rate and 98.85% True Negative (TN) rate, nearly matching performance on both classes. The model effectively learns specific intrusion patterns previously overlooked, leading to a drastic reduction in False Positives (FP) to 1.15%, enhancing operational efficiency in IIoT environments. This enhanced symmetry and high sensitivity underscore a robust, unbiased system capable of minimizing unnecessary, disruptive manual investigations.

Table 7 Confusion Matrix for ADASYN-Balanced Dataset

	Predicted: Normal	Predicted: Attack
Actual: Normal	TN: 98.85% (of Normal)	FP: 1.15% (of Normal)
Actual: Attack	FN: 1.60% (of Attack)	TP: 98.40% (of Attack)

The transition between these two matrices proves that the Computer Vision approach is highly capable, but it requires a balanced feature space to reach its full potential. Without ADASYN, the model is "lazy"—it guesses "Normal" most of the time to maintain high accuracy. With ADASYN, the model becomes "attentive," effectively distinguishing between subtle pixel-level differences in traffic images.

4. Conclusion

This research successfully demonstrates the efficacy of a Computer Vision-based Intrusion Detection System for IIoT networks. By converting network traffic into image-based representations and utilizing an ANN Binary Classifier, the system achieved high-fidelity detection capabilities. A critical finding of this study is the significant impact of data balancing; while the original imbalanced HiTar-2024 dataset yielded respectable results, the application of ADASYN propelled the model's Recall from 81.20% to 98.40% and reduced the False Positive Rate to 1.15%. This proves that addressing class imbalance is not merely an optimization but a necessity for developing reliable security frameworks in Industrial IoT environments, where missing a single intrusion (False Negative) can have catastrophic operational consequences.

6. Future Studies

Moving forward, several avenues exist to enhance the robustness and scalability of the proposed ANNBC system:

i. Multi-Class Classification: Transitioning from a binary "Attack vs. Normal" model to a multi-class classifier to identify specific types of IoT threats (e.g., DDoS, Sinkhole, or Man-in-the-Middle attacks) within the HiTar-2024 dataset.

ii. Deep Learning Architectures: Exploring the use of Convolutional Neural Networks (CNNs) or Vision Transformers (ViTs). Since the methodology involves transforming traffic into images, these architectures may extract more complex spatial features than a standard ANN.

iii. Real-Time Deployment: Testing the system's computational latency on edge devices (e.g., Raspberry Pi or Jetson Nano) to ensure the computer vision transformation and inference can occur in real-time without bottlenecking IIoT traffic.

iv. Hybrid Sampling Techniques: Comparing ADASYN with other hybrid methods like SMOTE-Tomek to further refine the decision boundary and minimize noise in the synthetically generated data.

References

1. Ahmed, S. F., Alam, M. S. B., Hoque, M., Lameesa, A., Afrin, S., Farah, T., . & Muyeen, S. M. (2023). Industrial Internet of Things enabled technologies, challenges, and future directions. *Computers and Electrical Engineering*, *110*, 108847.

2. Hu, Y., Jia, Q., Yao, Y., Lee, Y., Lee, M., Wang, C., . & Yu, F. R. (2024). Industrial internet of things intelligence empowering smart manufacturing: A literature review. *IEEE Internet of Things Journal*, *11*(11), 19143-19167.

3. Pradhan, A., Das, S., Piran, M. J., & Han, Z. (2024). A survey on physical layer security of ultra/hyper reliable low latency communication in 5G and 6G networks: Recent advancements, challenges, and future directions. *IEEE Access*, *12*, 112320-112353.

4. Andrade, R. O., Tello-Oquendo, L., & Ortiz, I. (2021). Cybersecurity Risks of IoT on Smart Cities. In *Cybersecurity Risk of IoT on Smart Cities* (pp. 1-22). Cham: Springer International Publishing.

5. Okoli, U. I., Obi, O. C., Adewusi, A. O., & Abrahams, T. O. (2024). Machine learning in cybersecurity: A review of threat detection and defense mechanisms. *World Journal of Advanced Research and Reviews*, *21*(1), 2286-2295.

6. Moustafa, N., Koroniotis, N., Keshk, M., Zomaya, A. Y., & Tari, Z. (2023). Explainable intrusion detection for cyber defences in the internet of things: Opportunities and solutions. *IEEE Communications Surveys & Tutorials*, *25*(3), 1775-1807.

7. Tsiknas, K., Taketzis, D., Demertzis, K., & Skianis, C. (2021). Cyber threats to industrial IoT: a survey on attacks and countermeasures. *IoT*, *2*(1), 163-186.

8. Vetrivel, S. C., Maheswari, R., & Saravanan, T. P. (2024). Industrial IOT: security threats and counter measures. In *Communication Technologies and Security Challenges in IoT: Present and Future* (pp. 403-425). Singapore: Springer Nature Singapore.

9. Almukhalfi, H., Noor, A., & Noor, T. H. (2024). Traffic management approaches using machine learning and deep learning techniques: A survey. *Engineering Applications of Artificial Intelligence*, *133*, 108147.

10. Wilson, A., & Anwar, M. R. (2024). The future of adaptive machine learning algorithms in high-dimensional data processing. *International Transactions on Artificial Intelligence*, *3*(1), 97-107.

11. Ma, B., Guo, W., & Zhang, J. (2020). A survey of online data-driven proactive 5G network optimisation using machine learning. *IEEE access*, *8*, 35606-35637.

12. Zhao, J., Masood, R., & Seneviratne, S. (2021). A review of computer vision methods in network security. *IEEE Communications Surveys & Tutorials*, *23*(3), 1838-1878.

13. Genuario, F., Santoro, G., Giliberti, M., Bello, S., Zazzera, E., & Impedovo, D. (2024). Machine learning-based methodologies for cyber-attacks and network traffic monitoring: A review and insights. *Information*, *15*(11), 741.

14. Ferrag, M. A., Friha, O., Hamouda, D., Maglaras, L., & Janicke, H. (2022). Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning. *IEEe Access*, *10*, 40281-40306.

15. Ahmad, B., Wu, Z., Huang, Y., & Rehman, S. U. (2024). Enhancing the security in IoT and IIoT networks: An intrusion detection scheme leveraging deep transfer learning. *Knowledge-Based Systems*, *305*, 112614.

16. Kotecha, K., Verma, R., Rao, P. V., Prasad, P., Mishra, V. K., Badal, T., . & Sharma, S. (2021). Enhanced network intrusion detection system. *Sensors*, *21*(23), 7835.

17. Halim, A. M., Dwifebri, M., & Nhita, F. (2023). Handling imbalanced data sets using SMOTE and ADASYN to improve classification performance of Ecoli data sets. *Building of informatics, technology and science (bits)*, *5*(1), 246-253.

18. Dhaouadi, T., Mrabet, H., Alhomoud, A., & Jemai, A. (2025). An intrusion detection system based on HiTar-2024 dataset generation from LOG files for smart industrial internet-of-things environment. *Computers, Materials, & Continua*, *82*(3), 4535.