Journal of Current Research and Studies

Volume 02 Issue 04 July-August 2025
https://journalcurrentresearch.com/
ISSN:3048-9911

OPEN ACCESS

***Correspondence**
Adeola Olajide Olatunde

Article Received
08/08/2025
Accepted
14/08/2025
Published
14/08/2025

Works Cited

Adeola, Olajide Olatunde; Alese, Boniface Kayode; Akinwonmi, Akintoba Emmanuel; Owolafe, Otasowie & Omoniyi, Victoria Ibiyemi. (2025). Detection of Real-Time Anomalies in Network Environment Using Deep Learning. Journal of Current Research and Studies, 2(4), 103-119.

Detection of Real-Time Anomalies in Network Environment Using Deep Learning

The Evolution of Malware Sandboxing and Artificial Intelligence for Smart Threat Detection

[1]Adeola, Olajide Olatunde; [2]Alese, Boniface Kayode; [1]Akinwonmi, Akintoba Emmanuel; [2]Owolafe, Otasowie & [3]Omoniyi, Victoria Ibiyemi

[1]Department of Computer Science, Federal University of Technology, Akure, Ondo State, Nigeria
[2]Department of Cyber Security, Federal University of Technology, Akure, Ondo State, Nigeria
[3]Department of Software Engineering, Federal University of Technology, Akure, Ondo State, **Nigeria**

*Corresponding Author: Adeola, Olajide Olatunde

Abstract

The exponentially growing overlap of the networks of Information Technology (IT) and Operational Technology (OT) of which the widespread establishment of the Internet of Things (IoT) is a main feature of Industry 4.0 has greatly increased the attack surface targeted at the critical infrastructure. This convergence is not only providing massive operational efficiencies; it is also creating a new form of unprecedented cyber-physical risks and this requires the deployment of advanced anomaly detection processes. The signature-based security tools have proved inefficient due to the increase in sophistication and invention of new cyber threats, analysis of zero-day attacks, and clandestine insider threats. Deep learning (DL) has become an efficient paradigm, and it is the only technology that can analyse the massive, multi-dimensional and high-dimensional data streams originating in a decentralised way as far as recognising in complex patterns and subtle deviations that may signal of malicious. This paper categorizes the algorithm based on deep learning that detects anomalies and provides an in-depth discussion of the recent deep learning-based anomaly detection algorithms with emphasis on the context that is related to these converged environments. It also evaluates critically the application of the various DL structures such as Autoencoders, Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), Transformers and Graph Neural Networks (GNNs) in the identification of the threatening activities, comprising Distributed Denial of Service (DDoS) attacks, malware as well as cyber-physical. Along with this, the paper illuminates the great difficulties that occur in practice (deployment), including but are not limited to data scarcity and imbalance, interpretability, computational overhead, and adversarial. Among the most prominent trends, one is aware of the transition of the models to unsupervised and hybrid with a special focus on the necessity to consider the explosion of Explainable AI (XAI), Federated Learning (FL) and the necessity of solid design of AI. The paper concludes by highlighting some of

the main areas where future research can be done to come up with more trustworthy, believable and practically usable anomaly detection system in strategic critical infrastructures.

Keywords

Anomaly, Deep Learning, Detection, Network, Models

1.0 Introduction

Background on Converged IT/OT Networks and the Evolving Threat Landscape

In the past, the Information Technology (IT) and Operational Technology (OT) networks were kept separate and independent spaces. Nevertheless, a fast-increasing process of convergence between the same two worlds characterizes the modern environment. It has been first and foremost due to digital transformation projects, emergence of Industry 4.0, and prevalence of device Internet of Things (IoT) in the industrial environment (Mishra and Pandya, 2021). Whilst CAS has a lot of potential in terms of improved data analytics and operational efficiencies, as well as a greater degree of automation, it also places unparalleled pressure on cybersecurity. Such a cross-over hugely increases the surface area that can be used to attack significantly introduces new vectors of increased cyber-physical assaults capable of leveraging the exposure of both IT and OT systems (Abshari, 2025). This interdependent nature means that these traditional security technologies which were intended to provide security in standalone IT or OT systems no longer suffice to give full protection. Such converged environments further make security efforts financially involving as they have quite striking features. These involve their natural heterogeneity of the devices and communication protocols, the especially tight real time operation requirements, and the high safety concerns and implications, which distinguish them fundamentally in comparison with traditional IT networks. What is at stake is the need to have a holistic and integrated anomaly detection method that can appreciate that there exist interdependencies among IT and OT and leave behind piecemeal security to achieve critical infrastructure resilience (Alparslan, 2023).

The Significance of Anomaly Detection (AD)

The vital cybersecurity system is called anomaly detection and is aimed at detecting any deviations to the accepted normal behavior within the system (Ododo and Addotey, 2025). It is significant because it helps to detect a wide range of threats in an early stage such as network intrusions, malware intrusions, insider and malfunctions of operations in industrial systems. The dynamism and constantly changing face of recent cyber threats are increasingly becoming a problem to traditional detection mechanism that is based on signature-based detection that relies on a database containing patterns relating to known threats. Such approaches have the most difficulty coping effectively with new, polymorphic, and zero-day attacks because they have no pre-defined signatures (Rashidi et al., 2025). This fundament limitation implies that adaptive, intelligent, and proactive strategies must be adopted because they are capable of detecting deviances in unknown threats not because of the specifics turn of events, but because of the behavioral patterns.

These constraints on the traditional signature-based strategy against emerging, complex and zero-day threats introduce a strong desire in the necessity of a new paradigm in cybersecurity. Where what is in place today in terms of security protection are not working against new and emerging attacks, a more dynamic and smarter one becomes critical. This fuels the implementation of AI-based anomaly detection which has the potential to detect so-called zero-day attacks via their exceptional behavior detection (Vinayakumar et al., 2019). It is also a paradigm-shift in cybersecurity because the presently known-threat reactive perception of adversarial behavior is being changed into a more proactive and behavioral perception of anomaly detection. The current trend is important in ensuring reliability and accessibility of present networked systems.

Role of Artificial Intelligence- Deep Learning

Artificial Intelligence (AI) and Deep Learning (DL) have proved to be the game-changers in the anomaly detection business (Alger and Tu, 2025). The DL models have one distinctive opportunity to work with huge, complicated, and multi-dimensional data received by different sources including the network traffic and industrial sensors (Fotiadou et al., 2020; Xu, 2025). They can learn complicated patterns and identify implicit abnormalities that cannot be identified with the help of the usual methods. One of the major strengths of DL is that this strategy allows the extraction of the relevant components of raw data itself, with no prior production of features, which significantly decreases manual work in the feature engineering endeavor (Alger and Tu, 2025). This feature makes it very appropriate to work with such large amounts of data and complexity in the data due to the heavily used modern computer networks and industrial systems.

The fact that conveys Deep Learning (DL) is effective is that it can learn complex, hierarchical features directly as opposed to raw, high-dimensional data because this is inherently the way humans think. This presents a source of costly and time-consuming hand-engineering of features on classic machine learning. However, unlike DL, its strength is represented in its ability to learn features automatically (i.e., representation learning) by mining the data without given features (Alger and Tu, 2025). This is essential considering the excessive, complicated IoT data and large amounts of network traffic in such new systems. This factor will help overcome the complexities and scale of such environments, and thus DL will be a useful and required tool to have in anomaly detection in such settings. Automated identification of patterns across a huge amount of information makes deep learning one of the most beneficial technologies to improve the cybersecurity of an interrelated world (Ibrahim et al., 2020).

The central goal of the paper is to carry out a critical review of current computer networks-and industry specific areas of anomaly detection techniques. The review attempts to learn the work on solving the particular issues with the use of the deep learning, the scores obtained, the restrictions determined by the methods used, and some general issues which still make the use of these techniques more problematic in the real world. The contribution of the paper is a synthesis of the existing body of knowledge, the identification of major unknown trends, the identification of major gaps in the knowledge, and suggestions on important future directions of creating more solid and reliable AI-powered anomaly detection systems in these critical infrastructures.

2. Literature Review

Evolution of Anomaly Detection in Networked Systems

The sphere of anomaly detection has been changed radically due to the constantly augmenting levels of advanced and the load of threats in cyber conditions. The first solutions were more signature-based systems that compare traffic (or look at system behavior) and compare it against a list of known signatures of an attack (Goetz and Humm, 2024). Given that a set such as this was effective in stopping previous threats known to the developers, these kinds of defenses are reactive, and thus have to be out of date when dealing with new, polymorphic, or zero-day attack (Alger and Tu, 2025). Statistical methods, appearing as a subsequent advancement, tried to provide a model of normal behavior of the system and raise an alarm when deviations missed a certain statistically defined boundary. Nevertheless, the current approaches tended to have unsatisfactory performance due to the dimensionality and complexity of modern data on the network, having a high level of false positives and little flexibility.

It was the drawbacks of these previous techniques, especially in their inability to keep pace with the increasing scale, complexity and peculiarities of the threats of the modern era of cyber-attacks that preconditioned the introduction of additional effectively working Artificial Intelligence and, in particular, deep learning techniques (Goetz and Humm, 2024). The evolution of the signature-based to AI-aid protocols spells out a surge in the trend towards the usage of more adaptive and intelligence machine reaction that is directly proportional to levels and complexity of cyber attacks. The simpler processes which sprung up earlier were overly structured and could not do much to counter more advanced and unfamiliar cyberattacks. This incompetency gave rise to the development of more intelligent systems on learning (Alger and Tu, 2025). The history of anomaly detection is therefore a direct reflection of the never-ending cat and mouse game that takes place in the world of cybersecurity where the constantly changing nature of attack

techniques has overpowered the one dimensional nature of constant defense. The change was organically transferred to the aspect of detection where it is necessary to develop adaptive and intelligent systems with the ability to identify previously unknown threats.

Architecture of Deep Learning for Anomaly Detection

Deep learning models have the advantage of using multi-layered neural networks that learn to recognize more complex patterns since learning hierarchical data representation is what allows them to be highly effective in regards to establishing anomaly detection (Wang et al., 2025). However, not every DL architecture is arbitrary and each of them is trained or tuned to address specific data qualities or problems and this trend is being directed to increasingly specific and amplitude optimized qualities (Wang et al., 2025). Time-series, image like, or graph structured data of many types will have different structure and thus require different strengths in the model to capture the true pattern of the data. It translates into the realisation that there is no one-to-one specific model that can be deemed as the best one, rather, when applied to particular types of anomaly detection, the set of specialised DL tools is needed.

The Autoencoders (AEs) and Variational Autoencoders (VAEs)

Autoencoders are the example programs of unsupervised neural networks that have been utilized a long time in anomaly detection, as they are able to acquire low dimensional features which portray normal data Wang et al., 2025). The basic concept is, that an AE will be trained to predict its input, with a modicum of inaccuracy. In the training, the network is taught with the pattern and structure of normal data in it. In case of an anomalous input, the presented input is far away in terms of error from the learned patterns and this is associated with a large reconstruction error. This large error acts as a great indication of an anomaly and therefore this would be easily discovered. They are useful in a wide range of data such as network traffic, system logs, and even data produced by industrial sensors and as such they are good at detecting minor anomalies in known normal data patterns (Wang et al., 2025). Variational Autoencoders (VAEs) are extensions of AEs that incorporate a stochastic element in order to enable such networks to generate new objects and interpretations of the training data.

Long Short-Time memory (LSTMs) and Recurrent Neuronal Networks (RNNs) and Gated Recurrent Unit (GRUs)

The RNNs and its more advanced variants like LSTMs and GRUs have been specifically developed to be applied to certain types of sequential and time-series data familiar to network traffic analysis and Industrial Control System (ICS) sensor data (Fotiadou et al., 2020; Jaouedi et al., 2020). They are also provided with the inner memory in which temporal dependability in the data streams is covered (Wang et al., 2025). This is an essential ability to recognizing anomalies that exhibit themselves across a span of time as an uncommon sequence or even pattern, e.g., Distributed Denial of Service (DDoS) attacks, which entail excess traffic at an unexpected rate over a time frame, or slight deviations of processes being undertaken in an industrial set-up that occur with the passage of time (Alger and Tu, 2025). The two sequential networks, LSTMs and GRU are especially efficient in learning of long-term dependence since the vanishing out (but to certain extent still present in LSTMs and GRU) has been mitigated.

Convolutional Neural Networks (CNNs)

CNNs have widely applied anomaly detection on data that is not visual, despite being conventionally recognized as such effective image processors (Fotiadou et al., 2020; Alger and Tu, 2025). With regard to local pattern extraction and spatial feature extraction of structured data, they operate superbly. CNNs have the potential to study network packet header information, system logs, or even transformed time-series data (e.g., by converting sequences of measurements to 2D representations such as spectrograms) in the network (Fotiadou et al., 2020). Their capacity to detect local patterns makes them useful in activities such as the detection of malware because binary code can be interpreted as an image to identify malicious patterns and even network intrusion detection through looking at patterns in network flow data (Rashid et al., 2025).

Generative Adversarial Networks (GANs)

GANs are a special anomaly detector in the sense that they can play two roles. There are two adversarial neural networks, the discriminator and the generator. The generator is trained to generate data that has a distribution resembling that of the normal data but the discriminator tries to learn how to differentiate between real and generated data (Dunmore et al., 2023). The discriminator part can be trained with training examples of different instances which are not part of this derived normal distribution to be deemed as anomalies in anomaly detection. More importantly, the generator part of GANs is also capable of generating authentic anomaly data. This will especially help in overcoming the issue of data imbalance and scarcity that are usually involved in anomaly detection datasets, which contain very few examples of the real anomalies since anomalies are not frequent and therefore the data that is available in the curated datasets is usually biased and imbalanced, but by using this the training will be able to create representative anomalous samples that can be used to train the system (Dunmore et al., 2023).

Graph Neural Networks (GNN)

GNNs are a rather recent and effective method of anomaly detection in the scenarios of complex networks. It is unique to neural networks in the sense that instead of working with points in Euclidean space, as is done in traditional neural networks, GNNs work with data as a graph. This enables them to identify anomalies efficiently which could take the form of structural change in the network, communication patterns between nodes and also unusual propagation patterns within the network graph (Birihanu and Lendák, 2025). An example is that a GNN can detect an abnormal host, such as to detect an unusual connectivity or dataflow pattern that is not part of the normal patterns of the graph.

Transformer Models

Transformers were initially built to work with the natural language processing, but they proved to be extremely successful in time-series anomaly detection. Their main innovation is the self-attention mechanism enabling them to give different areas of an input sequence different weigths when doing predictions. This grants them the ability to more effectively learn long-term dependence in the sequences than the typical RNNs, and thus they are good for complex time series data and discovering recondite, time-prolonged anomalies that models with shorter memory cells would overlook (Ghourabi, 2022). They also have a benefit in terms of parallelization of processing.

The table below summarizes these outstanding deep learning models and their features with regard to anomaly detection:

Table 1: Overview of Deep Learning Models for Anomaly Detection

Model Type	Core Principle	Strengths for Anomaly Detection	Typical Applications in Anomaly Detection	Key Limitations
Autoencoder (AE) / VAE	Unsupervised learning of data representation; high reconstruction error indicates anomaly.	Effective for unsupervised anomaly detection; learns normal data distribution; good for high-dimensional data.	Network traffic, sensor data, system logs, user behavior.	Can struggle with complex, varied anomalies; sensitive to hyperparameter tuning; "black-box" nature.
RNN / LSTM / GRU	Processes sequential data; captures temporal dependencies and patterns over time.	Excellent for time-series data; detects anomalies in sequential behaviors (e.g., unusual sequences of events).	DDoS detection, ICS process monitoring, user behavior analysis, network flow analysis.	Vanishing/exploding gradients (less so for LSTM/GRU); high computational cost for very long sequences; difficulty with parallelization.

Model Type	Core Principle	Strengths for Anomaly Detection	Typical Applications in Anomaly Detection	Key Limitations
CNN	Extracts local features and patterns through convolutional filters.	Good for spatial feature extraction; effective on grid-like data (e.g., packet headers, converted time-series).	Malware detection (binary images), network intrusion detection (packet/flow features), log analysis.	May not capture long-range temporal dependencies well; requires data to be structured spatially.
GAN	Generative model learns normal data distribution; discriminator identifies deviations.	Can generate synthetic anomaly data to address imbalance; discriminator effective at identifying out-of-distribution samples.	Network intrusion detection, industrial system anomaly generation, rare event detection.	Challenging to train (mode collapse, instability); sensitive to hyperparameter tuning; can be computationally intensive.
GNN	Models relationships between entities in graph structures; detects anomalies in network topology/interactions.	Ideal for relational data; identifies structural anomalies, abnormal communication patterns in complex networks.	Social network anomaly detection, network topology analysis, cyber-physical system interaction monitoring.	High computational cost for large graphs; data representation as graphs can be complex; limited interpretability.
Transformer	Uses self-attention mechanisms to weigh importance of different sequence parts; captures long-range dependencies.	Excellent for long-range temporal dependencies; highly parallelizable; robust feature learning from sequences.	Time-series anomaly detection (network traffic, sensor data), log analysis, system call sequences.	High computational cost and memory usage for very long sequences; requires large datasets for effective training; interpretability can be challenging.

Distributed Denial of Service (DDoS)

Distributed Denial of Service (DDoS) attacks are called to flood a target device or network resource, hence rendering it not accessible to other clients. The trait of DDoS attacks that makes it very dynamic and with high traffic would be a problem in regards to the detection using simple thresholding. Insider Threats Preventible LSTMs and GRUs are particularly great at identifying time-related features of the network traffic: this includes unusual packets bursts on the network or traffic volatility over the time (Jaouedi et al., 2020). CNNs, however, are able to derive spatial information in the network traffic, e.g., patterns in the header of the packets or flow statistics and, as such, are effective against even more advanced types of DDoS attacks that are executed by obfuscating the traffic or by spoofing legitimate traffic patterns (Fotiadou et al., 2020). They are very effective because of their capability to learn both the temporal and spatial aspects.

Malware Detection

The introduction of advanced and multi-polymorphic malware is a major threat to conventional signature-based techniques of detecting malware that fail to detect the new variants (Vasan et al., 2020). An effective alternative is available through deep learning that can analyze different sources of data on malware. It involves investigation of the network traffic caused by malicious programs, analysis of sequence of system calls invoked by suspicious programs, or even use of raw binary code as an input into CNNs (Rashid et al., 2025). Using an example, CNNs would learn the binary code of an image and a corresponding malicious pattern of structures would signify its maliciousness, whereas, a sequence of abnormal system calls would identify the RNNs (Rashid et al., 2025). These DL approaches could detect

malicious codes that they have not seen before and know the habits of healthy programs in order to detect unusual functioning of a program on their systems, a significant improvement compared with signature-based approaches.

Identification of a Cyber-attack in Industrial Systems (OT/ICS/ SCADA)

The deep learning has been given special use on the Operational Technology (OT) environment, Industrial Control Systems (ICS) and Supervisory Control and Data Acquisition (SCADA) in order to perform the task of anomaly detection (Goetz and Humm, 2024). That there is such a disparity and ubiquity of problems in OT environments because of data scarcity, hard real-time constraints, safety criticality, inherent heterogeneity, and the necessity of domain experts, even general solutions involving deep learning require a significant amount of adaptation or supplementation with specialized techniques. This gives rise to the insight that OT protection demands an exceedingly specific, not to mention frequently multifaceted, AI solution, and it can no longer depend upon ready-made products. OT environments are dramatically different than regular IT networks, and the latter are the direct cause of the increase of the general AI woes; researchers and practitioners have to look into specific solutions.

Cyber-Physical Attack Detection

AI can be highly useful in detecting orchestrated attacks that can involve both IT and OT assets in order to control the physical process (Abshari and Sridhar, 2025). The cyber-physical attacks, such as the one referred to as Stuxnet, are aimed at damaging or disrupting physical systems through breaching control systems. It is also possible that deep learning models can help to plot any anomalies identified on the IT large (e.g. suspicious connected activities, Malware circulation) with anomalies identified on the OT sensors (e.g. the unexpected valve position, unexpected pressure value) (Birihanu and Lendák, 2025). Depending on multiple streams of data and being jointly processed, DL models can detect sophisticated threats that combine a chain of activities both at the domain levels, which will not be detected by disparate detection computers (Alger and Tu, 2025). Such a heterogeneous correlation of this kind is a necessity of a broad cyber-physical security.

Detection of Process Deviation and Sensors Faults

Autoencoders and RNNs have been of great use to regulate malfunctions or sensor signal measurements in an industrial setting (Drewek-Ossowicka et al., 2020). Such anomalies may point toward some cyberattack that is trying to control the physical processes or failure within operations of the crucial equipment. To illustrate, an Autoencoder may be trained on the range and correlation of the sensor readings to the extent that a large reconstruction error results in it dubbing the known reading as a read anomalously. The RNNs, however, have the ability to detect atypical temporal behaviour in the sensor readings that indicate a change in a process (Yang et al., 2021). It is emphasized that the real-time and safety-critical characteristics of the corresponding detections should be highly rigorous; all delays or inaccuracies can lead to some significant physical impact within the scope of equipment, environment, or even human losses.

Operational Technology (OT) Environment Peculiarity

The practical application of deep learning in OT setting is full of challenges that are unique and of high complexity, which are not present to the same level and extent in their IT counterparts (Goetz and Humm, 2025).

Scarcity and Unbalanced Data

The datasets of bad labeled data records that are both big and contain anomalies are a very difficult acquisition to find in ICS/SCADA environments (Albanbay, 2025). It is problematic since the likelihood of actual event occurrences is rare in well-run operational systems and to produce synthetic anomaly data in an operational system is costly, hazardous and unethical to do so. Restrictions and exposures of safety are concurrent with one another Anomalies can be learned with respect to real-world instances but are rare; therefore, learning can be difficult when using a supervised model. Such gross imbalances of usual and weird data cause models biased to the majority, and poor performance of the rare and anomalous events (Albanbay et al., 2025).

Real-time Constraints and Safety Implications

In many cases, the Operational Technology systems have strict latency requirements in terms of anomaly detection (Goetz and Humm, 2025). During a lot of industrial processes, the need to take a lot of action in the shortest period is sought because it is one of the provisions to ensure safety, no physical damages are incurred, or minimal financial losses are incurred (Goetz and Humm, 2025). An anomaly may cause disastrous results even in case of millisecond delay in detecting it. Moreover, the harsh outcomes of a false positive or false negative in the industrial processes that are safety-critical deserve top priority. A false positive can cause an unwanted shut down and lead to losses of production, a false negative can escalate a full-blown malicious attack or other critical failure and can be expensive in terms of bodily harm or environmental destruction (Birihanu and Lendák, 2025).

Lack of Standardization and Heterogeneity

In most cases, the industrial setting is not homogenous amongst the different systems in the way the IT setting is with various systems being built up on proprietary communication protocols (e.g. Modbus, DNP3, OPC UA) and other hardware. Such a natural disparity makes the whole process of data integration and normalization, in particular, the process of training AI models particularly challenging (Goetz and Humm, 2025). The Matching of The Domain of Knowledge The task of coming up with a standard deep learning model that could cleanly scale along this diverse ecosystem is by far immature, and usually, requires an ad-hoc solution to each specific implementation. A new experimentation of this blend of deep expertise of OT engineers and operators is needed to accomplish the successful deployment of deep learning in OT systems (Goetz and Humm, 2025).

The Integration of Domain Experts

The deployment of deep learning in OT systems would only be successful when coupled with a new fusion of deep expertise available to OT engineers and operators (Birihanu and Lendák, 2025). Their experience is particularly critical in a number of areas: precise anomaly labeling (the difference between common operational fluctuations and real anomalies), efficient feature engineering (relevant parameters of the operation), and the proper interpretation of anomaly warnings provided by AI in the situation of a particular operation (Birihanu and Lendák, 2025). Such a domain-specific knowledge may allow an AI model to avoid double-counting of benign changes in operations, as well as not labeling context-sensitive, small anomalies, which can cause false alarm rates to become excessive or cause the omission of detections.

This is due to the fact that the nature of OT environments presents extremely unique and omnipresent problems, where most broad-based deep learning will have to be heavily adapted or augmented by unique considerations or approaches, such as Transfer Learning or Hybrid Models. This brings about the realization that OT security involves a very sophisticated and, in many cases, a multi-pronged AI solution that is way beyond off-the-shelf solutions.

Table 2: Comparative Analysis of Anomaly Detection Datasets

S/N	Dataset Name	Primary Domain	Data Type	Specific Anomaly Types Covered	Labeled/Unlabeled Status	Characteristics/Challenges
1	CIC-IDS2017 / CIC-DDoS2019	IT Network	Network Flow/Packet	DDoS, Brute Force, Web Attack, Botnet, Infiltration	Labeled	Publicly available, widely used, but may not reflect latest attack methods.

2	CSE-CIC-IDS2018	IT Network	Network Flow/Packet	Brute Force, DDoS, DoS, Web Attack, Botnet, Infiltration, PortScan	Labeled	More diverse attacks than CIC-IDS2017, but still synthetic.
3	UNSW-NB15	IT Network	Network Flow/Packet	Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms	Labeled	Synthetic, older, but still a benchmark for various attacks.
4	WADI / SWaT	OT/ICS	Sensor Readings, Network Traffic	Cyber-physical attacks, process deviations	Labeled (some attacks)	Real-world ICS testbeds, but limited attack scenarios and data volume.
5	BATADAL	OT/ICS (Water Treatment)	Sensor Readings	Cyber-physical attacks	Labeled	Focus on water treatment, specific attack types.
6	IoT-23	IoT/IT Network	Network Traffic	DDoS, Mirai, Brute Force, XSS, DoS, Benign	Labeled	Focus on IoT device traffic, real-world captures.
7	LITNET-2020	IT Network	Network Flow	DDoS, Botnet, Brute Force, Port Scan	Labeled	Focus on modern network attacks, publicly available.
8	Edge-IIoTset	IoT/IIoT	Network Traffic, Logs	DoS, DDoS, MITM, Ransomware, XSS, SQL Injection	Labeled	Focus on edge IoT/IIoT, diverse attack types.
9	TON_IoT	IoT/IIoT	Network Traffic, Telemetry, Logs	DoS, DDoS, Ransomware, XSS, SQL Injection, Backdoor, Malware	Labeled	Comprehensive IoT/IIoT dataset, includes both network and telemetry.

10	General Challenges	Data scarcity for real-world anomalies, high imbalance, lack of standardization, proprietary nature of OT data.

Hybrid and Ensemble Deep Learning Models

The reasoning behind the implementation of several AI methods in order to detect anomalies is based on the fact that there is no single deep learning architecture that can be used to optimize the process of anomaly detection across the board. Rather, by taking advantage of the synergistic capabilities of various methods, it is possible to greatly increase the precision of the results, greatly minimize the false alarms, and achieve much more stability across virtually all forms of anomalies (Ahmed and Jameel, 2022). The latest trend, which implies the usage of hybrid and ensemble-based solutions (Jaouedi et al., 2020; Rashid et al., 2025), helps to capture such expediency. Because each individual deep learning model has its own strengths and weaknesses, it only stands to reason that the next step would then be to grow together so as to reduce weaknesses and increase strengths. This brings the creation of hybrid models, which directly translates to the enhanced level of accuracy in detecting results, and minimized false positives. This means a big and useful implication to practical applications where it is vital to devise a system to be robust and highly performing in a variety of conditions.

In the hybrid models, the implementation of other deep learning models is prevalent. To demonstrate an example, one can assume that a model involving CNN and LSTM architectures will be applied to find both spatial patterns (using CNNs) and time dependencies (using LSTMs) in network traffic data, which will result in the more detailed description of the elaborate patterns of the attackers (Fotiadou et al., 2020; Jaouedi et al., 2020). The other very typical way is the combination of deep learning and traditional machine learning algorithms. Here, deep learning models implicitly treat the task of features extraction as a robust feature learning, and they have the ability to automatically learn optimal representation of raw data that aids in final classification with conventional classifiers like Support Vector Machines (SVMs) or Random Forests (Schmitt, 2024). The reason is that it may benefit the strong aspect of feature learning potential of DL to preclude in accordance with the explainability or efficiency of traditional ML algorithms. Moreover, it is possible to transform saturation-aware improvements with deep learning (DL) and integration into statistical methods to learn the preliminary deviations and the fine-grained anomaly scoring or thresholding, respectively. The combination of these methods is specifically efficient in approach towards management of diversity and complexity of anomalies within converged IT/OT environments (Alger and Tu, 2025).

The Deep Learning Anomaly Detection, its flaws, and its disadvantages

Despite potential transformational properties, the use of the latter on anomaly detection in real-world PCs, industrial systems, and computer networks are constrained by some disadvantages and problematic areas (Schmitt, 2024). The fact that most of these issues come up in most studies points out to the fact that then those are not isolated problems but rather systematic problems in the application of AI in complex and dynamic security environments. The potential of DL is enormous, but at the same time, its most frequent usage has a number of weaknesses that require targeted research and engineering efforts. This brings about the realization that AI when applied to security is not a field that has been solved but is one under ongoing research and development and thus a continuous effort to realise the academic potential and transfer the same to real life application.

There are unbalanced Data and a shortage of Dataset.

Data imbalance on a massive scale might be seen as the big challenge. By their nature, anomalies are an infrequent occurrence when compared to the normal occurrence, and thus such models are biased towards the majority and perform poorly on the precise exceptional events (Sabeel et al., 2021). In the case that anomaly datasets are very skewed, a model can be very accurate by failing to classify most things anomalous. This drawback, although numerically high, is practically not very useful when it comes to detecting anomalies. This necessitates the use of assessment metrics that draw specific attention to the performance of the minority (anomaly) classes such as recall and precision (Vanin et al., 2022). Such can be redressed through numerous methods that involve oversampling minorities (e.g. SMOTE), down sampling majority populations, and manipulating synthetic information, particularly anomalies, with the use of Generative Adversarial Networks (GANs). With the artificial data the problem is that realistic and representative synthetics modelling the complex cyber-physical attack are hard to come by.

Explainable Artificial intelligence

The ASL as what is termed in the context of complex deep learning models poses a huge problem to trusting, adoption, and reliable debugging, especially those in critical and safety-critical systems such as an ICS (Sauka et al., 2023). On an occasion when an AI system highlights an abnormality, a security analyst and operators in the OT must be able to know why the decision has been given to make a case to confirm the alert, intervene appropriately, and eventually train on the event. The absence of interpretability will also imply that operators will possibly not be able to believe the system, and act effectively. The necessity of sound Explainable AI (XAI) methods to have transparency, explanation of model decisions, and provision of actionable insights to human operators and security analysts is now urgent (Birihanu & Lendák, 2025). This will enable the investigation to have a better insight into the cause of anomalies and contributes to improved incident response.

Just in time Processing and Computational Overhead

The computational necessities of training and inferences on model architectures of complex Deep learning models are challenging and present significant issues when it comes to real-time detection in fast networks and deployment at edge environments which may have limited computing capabilities (Wang et al., 2025). A network processing gigabits per second of network traffic or thousands of sensor readings per second is a huge level of processing power. The possible solutions to this would be optimizations of such models like quantization (weakening precision of connections), pruning (cutting unnecessary connections), and knowledge distillation (transference of information between a wide and a narrow net). Also, Edge AI placement that puts the computations closer to the data sources can offset the issues with latency and reduce the bandwidth utilization (Wang et al., 2025).

Adversarial Attack and Robustness

It has been demonstrated that the deep learning models are especially susceptible to small (and possibly imperceptible), pernicious modifications (adversarial attacks) to the input data (Alparslan, 2023; Sauka et al., 2023). These attacks may result in the misclassification of data, where a particular algorithm correctly classifies an abnormality or, on the contrary, leads to false detection, which can compromise the effectiveness and credibility of the security application based on AI (Alparslan, 2023). An example of this bypass attack entails the attacker making changes on an attribute of a network packet so that this change does not allow the intrusion detection system to recognize such a network packet as an illegal one. To mitigate this vulnerability, it should conduct critical studies on effective AI design as well as development of more effective adversarial training techniques to come up with more robust models that could not be easily manipulated by such advanced strategies and still retain their integrity within the hostile environment (Elgarhy et al., 2024).

Ethical Considerations

In addition to the technical concerns, the broader ethical risks of implementing AI in security include control of the law enforcement officials, algorithmic bias, and so on. This may result in discriminative or even an outright offered misclassification of correct actions as it can be caused by the fact that there exist biases in training data. Privacy issues

are also paramount when considering the processing of sensitive network traffic or user behavior which need to be taken into serious consideration when it comes to data anonymization and privacy-preserving mechanisms (Wang et al., 2025). Moreover, accountability of AI decision in the critical infrastructure is a critical point that should be considered; there must be frameworks to dictate who will bear responsibility when an AI system makes a wrong decision, with repercussions that are of high impacts.

3. Findings

Composition of the Adequate Deep Learning Approaches Relating to the Kind of the Anomaly and the Environment

The detailed literature review shows that the process of properly defining the best deep learning architecture and anomaly detection methodologies has produced a moderate level of understanding of best and worst approach to carrying this out with the efficacy dependent upon the anomaly detection task at hand and the network in play (IT or OT) (Xu et al., 2025; Wang, et al., 2025). The fact that certain deep learning architectures are much more effective at addressing a specific anomaly type implies a certain new theme of specialization and customization of the deep learning solution to the nature of the anomaly and the property of the data, given up on a one-size-fits-all approach. The fact that various models perform better at specific tasks of anomaly detection is logically consistent with the assumption that in case of various problems it is necessary to choose the most suitable model. This demands a more subtle appreciation of the positive and negative attributes of both models, and, as a result, one must conclude that successful anomaly detection applications in practice will require multiple dedicated deep learning modules operating cooperatively, rather than depending on a single general-purpose model.

LSTMs and GRUs have been widely successful in detecting temporal anomalies as the patterns in causes of anomalies often occur over a long time frame (Fotiadou et al., 2020; Jaouedi et al., 2020). On the other hand, CNNs are very useful when performing a task that requires the extraction of spatial features whether in solving the malware detection problem that uses raw binary code as a basis of analysis, or on tasks that need to extract specific patterns on network packet header (Rashid et al., 2025). GNNs have been gaining prominence as a potentially promising tool to identify peculiarities that occur as difficulties in the structure or distinct communication patterns in the complicated network topology. Autoencoders, thanks to their ability to learn independently, are becoming popular in terms of providing a benchmark anomaly detection task under different data types and their model tends to be popular when labeled anomaly data is limited (Wang et al., 2025).

Noteworthy is the growing popularity of hybrid models in attaining better performance through the meticulous integration of the advantages of different methods (Rashid et al., 2025). Combining CNNs with LSTMs, e.g., is another example of extending the networks to capture spatial as well as temporal features of network traffic to get a more complete picture of highly sophisticated attacks (Fotiadou et al., 2020). Similarly, the ensemble of deep learning architecture and classic machine learning classifiers to realize robust and in some cases interpretable feature extraction is possible to achieve better accuracy. These types of synergies utilize the power of many architectures to counteract one weakness to establish more dependable combinations of growing accurate recognitions of anomalies in diverse threat territories of converged IT/OT worlds (Xu et al., 2025).

New Paradigms and Trends

There are a few general tendencies and new paradigms influential in the domain of AI-based anomaly detection as the sphere continuously strives to address the limitations and expand functions of the systems. The emphasis on Explainable AI (XAI), Federated Learning (FL), and adversarial robustness as the key areas of future work, as well as the identification thereof as challenges currently of importance, demonstrates a definite connection: the latent weaknesses and insecurities of the existing AI models are creating the need to develop their trustworthiness, privacy-preserving abilities, and resiliency to well-organized attacks (Uccello et al., 2025). In the case of the current AI systems that have provable foundational issues e.g., they are often corruption resistant, having issues of being a black-box system, where data privacy is of concern and the ability to be attacked by adversaries, etc., future work should focus

on resolving these issues. This is what will result in the surfacing of these particular areas of research in terms of major themes, the transition taking place in the way AI development has been done so far to one that focuses more on ethical, trustful, and generally deployable research in AI systems where critical security applications play a vital role.

Among the most notable trends, it is possible to note the usage of unsupervised and semi-supervised learning models (Fotiadou et al., 2020). It is driven by the fact that shortage in data, particularly in professionally-labeled anomaly data, is still not solved and that new attacks or zero-day attacks require the capability to identify these anomalies even before any respective data have been marked. Unsupervised models are by design able to detect the unknown threats since they learn what system behavior is considered normal and any abnormal behavior is marked as suspicious.

The other major paradigm is the growing significance of the Explainable Artificial Intelligence (XAI). Since deep learning models are growing in complexity, the black box nature of their performance may be defined as one of the most significant barriers to creating trust, and thus is subject to implementation, particularly in safety-critical ICS, where the explainability of AI choices should become a priority (Sauka et al., 2023). The XAI methods are under development to make them transparent, explainable, and to provide actionable information to the human operators to induce higher confidence and promote successful incident response (Uccello et al., 2025).

Due to the existence of distributed data sources and issues concerning the privacy of anomaly-detection, privacy-preserving AI, and particularly, FL is becoming a viable prospect. FL allows training a model in a collaborative framework across two or more organizations, or divisible nodes on networks, without necessarily exchanging the input marks that are sensitive and restrict the capability to leverage diverse data (Xu et al., 2024).

Furthermore, much focus is given to the adversarial robustness. As we are aware, the adversarial attack on the deep learning models makes a loss in classification (Alparslan, 2023). Study is now going towards the implementation of AI models that are hardened against this kind of manipulation in order to guarantee their integrity and competence under experimentations by adversaries that have a deep interest in devising ways to disorient the security apparatus (Sauka et al., 2023).

Evaluation Methods and Performance Metrics

A critical review of the literature shows that there are severe issues involved in testing anomaly detection frameworks, especially in a scenario where a shaft has a very skewed dataset. The existential problem of data imbalance is the direct result of which a more sophisticated and suitable set of evaluation metrics rather than simple accuracy becomes necessary. This leads to an underlying trend of possible false performance implications when metrics are not thoughtfully selected in order to represent the actual effectiveness of the anomaly detection system within the rare, critical anomalous class. When anomaly datasets are highly skewed, the model can easily reach high-accuracy simply because it classifies the vast majority as something being normal, which is an almost useless result when it comes to anomaly detection. This will require evaluation measures that will critically evaluate this performance on the minority (anomaly) class. It means that the development of rigorous and appropriate assessment, performed by a wide array of metrics is as important as the development of a model, so the utility of anomaly detection tools and their reliability in real life could be ensured.

In most studies, there has been too much emphasis on basic accuracy as a core performance measure. Although accuracy is indicative of general accuracy, it is very misleading in imbalanced data sets where there are a lot of normal data in comparison to abnormal cases. A model that works well on identification of a normal case by 99% and does not identify any anomalous case may appear as an accuracy of 99% but it will not be useful in practice at all. Therefore, the awareness of the utter importance of such more clandestine measures as precision, recall (sensitivity), F1-score, and Receiver Operating Characteristic Area Under the Curve (ROC-AUC) has been on the rise (Alparslan, 2023; Birihanu, E. & Lendák., 2025). Precision is the fraction of the number of true positive prediction; recall is the reflection of the positive prediction that is the number of true positive prediction as a fraction of all the actual positive prediction. F1-score is in fact the harmonic mean between the precision and the recall that is susceptible to tradeoff. OC-AUC provides an evaluation of the intended correspondence between the true negative rate and the false negative rate as a number in different thresholds.

It is also important to take close care of false positive (Type I error) as well as false negative (Type II error) rates especially in important systems where the results of any misclassification could be disastrous (Ododo and Addotey, 2025). A high false positive rate may cause alert fatigue or other issues that may waste limited resources, and a high false negative rate may cause untenable consequences, since it will consequently overlook attacks that could result in serious loss.

Along with that, the published evidence has emphasized the high demand of standard datasets and reproducible evaluation benchmarks to make sure that results of the research are comparable and valid irrespective of the research (Wang et al., 2025). There are a lot of studies that utilise proprietary data sets or introduce an inconsistent preprocessing, which makes it hard to compare them. Further, the publicly available and labelled and representative datasets that reflect real IT/OT environments specifically, or types of anomalies in particular, and more general variations of normal traffic are also essential to the domain.

Significant Research Gaps and Unresolved Challenges

Nonetheless, the extensive array of research gaps and the yet unsolved issues remain impediments to the universal and trustworthy application of deep learning in anomaly detection in critical infrastructure despite significant progress made in this area. These challenges are not one-offs and therefore this is an indication that there are systemic obstacles to broad AI deployments in critical infrastructure. This implies that the current research needs a trend in which research that needs to be done in future must therefore focus more on these issues as being fundamental through which progress might really be achieved. When such basic issues are recurrently indicated in varying studies, this is a sign that they are not just passing problems but have been inherent problems which have not yet been solved. This makes them seen as key research gaps where a lot of focus is required meaning that even though, a substantial progress has been achieved, more fundamental work still needs to be done to ensure that AI is actually reliable, trustworthy, and broadly usable in sensitive and critical applications.

The data issue in terms of its persistence and complexity is the key one. This includes data collection issues, consistent labeling ensuring high quality of annotations, effective addressing of extreme data imbalance and heterogeneous data integrating data, network flows, sensor readings, system logs, security events, etc. Although it looks promising in regards to its ability to generate new data through GANs, the quality of these generation is still subject to improvement, particularly when it comes to complex cyber-physical environments.

Another such gap is that there are currently no robust and widely applicable explainable Explainable AI (XAI) techniques to efficiently tell why a complex and deep learning model thinks the way it does (Sauka et al., 2023). Alternative post-hoc explanation techniques exist; however, real-time detection of anomalies via a deep learning architecture that is inherently interpretable or an explanation that can be acted upon and performed in the context of the operation of a dynamic IT/OT environment is an open question. This is particularly crucial in matters concerning the trust that the human operators have on safety-sensitive systems.

Moreover, the possibility that the certified robustness of a system against the strong adversarial attacks could be very difficult to obtain is also a rather serious concern (Alparslan, 2023). Contrary to positive benefits to such strategies of adversarial training the resilience can be increased but often in a trade-off with performance of models or their generalization. On the one hand, the issue of coming up with provably robust deep learning models that can resist the vast amount of adversarial manipulation in addition to having a high level of detection accuracy is still under development.

Gaps in the research also pertain to seamless cross-domain anomaly detection. The integration and correlation of both the IT setting and OT setting with significantly different protocols, data representation, and delays is likewise not a straightforward one (Birihanu and Lendák, 2025). The current approaches will employ manual feature engineering, or generate naive data fusion, therefore, are less generalizable and scalable. The data presents the necessity of multi-modal deep learning architectures enabled to perform intrinsically integrative and learning streams of heterogeneous IT/OT data.

Finally, the attempt to build genuinely proactive, adaptive responding to anomalies strategies based on superior technologies like Deep Reinforcement Learning (DRL) is only a little bit open, nonetheless. Although AI has the aptitude to identify anomalies, autonomy in smart and safe responses in real world, particularly security-sensitive OT, there must be massive breakthroughs in autonomous decision-making, risk imperativeness, and human/AI collaboration.

Deployment Plans and Possible Challenges

Along the way to the working realization of anomaly detections systems based on AI there are practically incredible obstacles and issues. The implementation difficulties of practical deployment like the integration with the current systems, learning in the long term, and the human in the loop requirement point towards the inability of purely technical model performance to succeed in practice. This means that the effectiveness and use of AI-based anomaly detection depend on the operations, human, and system reasons in the end. An AI model with all the technical aspects present and in perfect order is rather useless in case it cannot be easily applied to the existing operational settings, cannot adapt itself to the changing conditions, or make a human operator unable to adequately communicate and trust their results. That is an issue which cannot insist on the practical issues beyond an extend of model goodness. This means that research and development should pay more attention to the lifecycle of AI systems, including the period of development all the way to deployment, maintenance, and interaction with humans, so that they are of practical use and are successfully adopted.

A significant complication lies in the fact that such new AI solutions have to be introduced into an already existing, frequently legacy, IT/OT infrastructure. A high percentage of industrial systems were not initially developed to consider the implementation of cybersecurity or AI in the system, and they possess the proprietary nature or criticalness of their operations, which makes any modifications extremely cumbersome and dangerous. This necessitates the use of strong APIs, protocol converters and consideration of the architectural design of the network so that the current operations are not rippled.

The second practical aspect is an urgent need to engage in lifelong learning and getting ready to drift in concepts. It operates in dynamic environments where behaviors of networks, system configurations, and patterns of attacks are changing all the time. A machine learning model that is taught to recognize anomalies using only historical data that will be stagnant will quickly become obsolete. To sustain effectiveness in responding to new threats and evolutions in normal behavior, operational systems need mechanisms of continuous monitoring, retraining and sometimes update of models.

Also, the human-in-the-loop factor is the most important. The AI-facilitated automatic detection of anomalies is not supposed to be a vacuum. The warnings by the AI should be accepted by human historian and OT operators and action taken and the AI. It is more of a note on the necessity of human-AI collaboration and how AI can enhance human capabilities as opposed to people replacing them. The user interfaces need to be obvious, alerts should be responsive, and the system should create adequate contexts and explanations so as to empower the human beings to make their decisions. The balancing of the automation and human monitoring in the safety systems of industrial products must be placed with concern with human interface design and policy.

5. Conclusion and Recommendations

The trend toward merging the Information Technology (IT) networks and Operational Technology (OT) networks has essentially transformed the former through the occurrence of difficult-to-detect, cyber-physical risks that are defined by prerequisites to master anomaly detection abilities. The present review will outline how deep into this transformation of the threat environment, deep learning can be a game-changer. Some of the deep learning architectures which displayed a significant level of success in grappling with the large, complex, of temporally-varying nature data stream which is characteristic of present networks are autoencoders, Recurrent Neural Networks, Convolutional Neural Networks, Generative Adversarial Networks, Graph Neural Networks, and Transformers. They also excel at automatic extraction of complex patterns and early identification of small deviations that have the potential of malicious behavior and surpass signature-based methods far beyond.

However, this path on the way to completely autonomous, trustworthy and resilient AI-based security solutions is not finished yet. Issues mentioned in the review have constantly included the existence of long-standing challenges, especially those surrounding data-data scarcity of labeled anomalies, extreme imbalances, heterogeneity in IT/OT data sources. Moreover, the fact that complex deep learning models are actually considered a black-box creates the problem of interpretability greatly limiting the possibilities of trust in and successful human intervention, particularly in industrial systems that may have a significant safety issue. The occurrence of the severe doubts regarding the vulnerability of the same models to advanced adversarial attacks are also questionable since the security and safety of the AI-based security systems can be questioned (Alparslan, 2023; Sauka et al., 2023).

On the basis of this review with the gap identified, it is possible to recommend several essential future research directions that can develop the field of anomaly detection in converged IT/OT environment such as Explained AI (XAI) based Anomaly Detection, Federated Learning (FL), Strong AI against Adversarial Manipulations, Hybrid and Multi-modal, Data Augmentation and Synthetic Data Generation.

References

1) Abshari, D. & Sridhar, M. (2025). A Survey of Anomaly Detection in Cyber-Physical Systems (2025). arXiv. Available at: https://arxiv.org/html/2502.13256v1
2) Ahmed, W. F., & Jameel, N. G. M. (2022). Malicious URL detection using decision tree-based lexical feature selection and multilayer perceptron model. UHD Journal of Science and Technology, 6(2), 105-116. https://doi.org/10.21928/uhdjst.v6n2y2022.pp105-116
3) Albanbay, N., Tursynbek, Y., Graffi, K., Uskenbayeva, R., Kalpeyeva, Z., Abilkaiyr, Z. & Ayapov, Y. (2025). Federated Learning-Based Intrusion Detection in IoT Networks: Performance Evaluation and Data Scaling Study. J. Sens. Actuator Netw. 2025, 14(4), 78. Available at: https://www.mdpi.com/2224-2708/14/4/78
4) Alger, J & Tu, M. (2025). Anomaly Detection of Network Layer Attacks Against Cyber Physical Systems Using Machine Learning and Deep Learning Techniques. Journal of Military Cyber Affairs, 8(1). Available at: https://digitalcommons.usf.edu/mca/vol8/iss1/4/
5) Alparslan, Y. C. (2023). Adversarial Attacks and Robustness in Deep Learning Models and Applications. Master of Science (M.S.), Drexel University Research Discovery. Available at: https://researchdiscovery.drexel.edu/esploro/outputs/graduate/Adversarial-Attacks-and-Robustness-in-Deep/991014961449004721
6) Birihanu, E. & Lendák, I. (2025). Explainable correlation-based anomaly detection for Industrial Control Systems. Front. Artif. Intell., Sec. Machine Learning and Artificial Intelligence. Available at: https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1508821/full
7) Drewek-Ossowicka, A., Pietrołaj, M., & Rumiński, J. (2020). A survey of neural networks usage for intrusion detection systems. Journal of Ambient Intelligence and Humanized Computing, 12, 497–514. https://doi.org/10.1007/s12652-020-02014-x
8) Dunmore, A., Jang-Jaccard, J., Sabrina, F., & Kwak, J. (2023). A comprehensive survey of generative adversarial networks (GANs) in cybersecurity intrusion detection. IEEE Access, 11, 76071-76095. https://doi.org/10.1109/ACCESS.2023.3296707
9) Elgarhy, I., Badr, M. M., Mahmoud, M. Alsabaan, M., Alshawi, T. & Alsaqhan, M. (2024). XAI-Based Accurate Anomaly Detector That Is Robust Against Black-Box Evasion Attacks for the Smart Grid. Applied Sciences, 14(21), 9897. Available at: https://www.mdpi.com/2076-3417/14/21/9897
10) Fotiadou, K., Velivassaki, T. H., Voulkidis, A., Skias, D., Tsekeridou, S., & Zahariadis, T. (2020). Network Traffic Anomaly Detection via Deep Learning. DOAJ. 12(5), 215. Available at: https://doaj.org/article/55d3922e9afe43619db7d1788842f561
11) Ghourabi, A. (2022). A security model based on LightGBM and Transformer to protect healthcare systems from cyberattacks. IEEE Access, 10, 48890-48905. https://doi.org/10.1109/ACCESS.2022.3172432
12) Goetz, C. & Humm, B. G. (2024). A Hybrid and Modular Integration Concept for Anomaly Detection in Industrial Control Systems. AI 2025, 6(5), 91. Available at: https://www.mdpi.com/2673-2688/6/5/91

13) Ibrahim, A., Thiruvady, D., Schneider, J. G., & Abdelrazek, M. (2020). The Challenges of Leveraging Threat Intelligence to Stop Data Breaches. DOAJ. Available at: https://doaj.org/article/6f5f971cbbb34aa4a892a7257b2ab8af
14) Jaouedi, N., Boujnah, N., & Bouhlel, M. S. (2020). A new hybrid deep learning model for human action recognition. DOAJ. 32(4), 447 – 453. Available at: https://doaj.org/article/06cbf63fdc83418f9d356cfe3c1b761e
15) Liu, H. & Lang, B. (2020). Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. (2020). DOAJ. 9(20), 4396. Available at: https://doaj.org/article/ecac058e11464fb2a61b606bc09d2b6e
16) Mishra, N., & Pandya, S. (2021). Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review. IEEE Access, 9, 59353–59386. https://doi.org/10.1109/ACCESS.2021.3073408
17) Ododo, F. & Addotey, N. (2025). Advancements And Challenges In Deep Learning For Cyber Threat Detection. ResearchGate. International Journal of Science Research and Technology. 7(9). Available at: DOI: 10.70382/tijsrat.v07i9.019.
18) Rashid, M. U., Khan, M. A., Alhaisoni, M., Tariq, U., Armghan, A., Alenezi, F., & Alqahtani, A. (2025). Hybrid Android Malware Detection and Classification Using Deep Neural Networks. International Journal of Computational Intelligence Systems, 18(1), 1–26. Avaialble at: https://doaj.org/article/85b61489ab0b4420917e50892a902364
19) Sabeel, U., Heydari, S. S., Elgazzar, K., & El-Khatib, K. (2021). Building an intrusion detection system to detect atypical cyberattack flows. IEEE Access, 9, 94352–94364. https://doi.org/10.1109/ACCESS.2021.3093830
20) Sauka, K., Shin, G. Y., Kim, D. W., & Han, M.-M. (2023). Adversarial Robust and Explainable Network Intrusion Detection Systems Based on Deep Learning. DOAJ. Available at: https://doaj.org/article/d11145a82960455084cc517e10227a57
21) Schmitt, M. (2023). Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection. Journal of Industrial Information Integration, 100520. https://doi.org/10.1016/j.jii.2023.100520
22) Uccello, F., Pawlicki, M., D'Antonio,S., Kozik, R. & Choraś, M. A. (2025). New Cybersecurity Approach Enhanced by XAI-Derived Rules to Improve Network Intrusion Detection and SIEM. Computers, Materials & Continua, 83(2), 2023–2045. Available at: https://www.techscience.com/cmc/v83n2/60589/html
23) Vanin, P., Newe, T., Dhirani, L. L., O'Connell, E., O'Shea, D., Lee, B., & Rao, M. (2022). A study of network intrusion detection systems using artificial intelligence/machine learning. Applied Sciences, 12(22), 11752. https://doi.org/10.3390/app122211752
24) Vasan, D., Alazab, M., Venkatraman, S., Akram, J., & Qin, Z. (2020). MTHAEL: Cross-architecture IoT malware detection based on neural network advanced ensemble learning. IEEE Transactions on Computers. https://doi.org/10.1109/TC.2020.3015584
25) Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. IEEE Access, 7, 101179-101190. https://doi.org/10.1109/ACCESS.2019.2906934
26) Wang, F. Jiang, Y., Zhang, R., Wei, A., Xie, J. & Pang, X. (2025). A Survey of Deep Anomaly Detection in Multivariate Time Series: Taxonomy, Applications, and Directions (MTSAD). Sensors 2025, 25(1). Available at: https://www.mdpi.com/1424-8220/25/1/190
27) Xu, Z., Wu, Y., Wang, S., Gao, J., Qiu, T., Wang, Z., Wan, H.& Zhao, X. (2025). Deep Learning-based Intrusion Detection Systems: A Survey. Cryptography and Security. Available at: https://arxiv.org/abs/2504.07839
28) Yang, J., Li, T., Liang, G., He, W., & Zhao, Y. (2019). A simple recurrent unit model based intrusion detection system with DCGAN. IEEE Access, 7, 83286-83296. https://doi.org/10.1109/ACCESS.2019.2922692