An overview of Detecting Attacks for Cyber security in Industrial Internet of Things: Literature Review

Maithili Shailesh Andhare
Assistant Professor, E&TC Engineering
PCET's PCCOE&R, Ravet
Pune, India
maithili.andhare@pccoer.in

Vijayalaxmi Sandeep Kumbhar
Assistant Professor, E&TC Engineering
PCET's PCCOE&R, Ravet
Pune, India
vijayalaxmi.kumbar@pccoer.in

Arti Avinash Tekade
Assistant Professor, E&TC Engineering
PCET's PCCOE&R, Ravet
Pune, India
arti.tekade@pccoer.in

***Abstract*— Cybercriminals and hackers are actively pursuing critical city infrastructures that rely on smart "Industrial Internet of Things (IIoT)" devices. Regardless of the fact that it has prompted a number of interests in recent decades, there isn't an accurate approach for Industrial IoT attack detection. Prior to actually developing an appropriate approach for detecting Industrial IoT attacks, it's indeed necessary to have knowledge of previous literature works. As a result, a concise and conceptual literature evaluation is conducted in this research work, including the most applicable methodologies dedicated to IIoT attack detection. All of the research papers gathered is from the years 2020 to 2022. Furthermore, each of the gathered publications is examined in terms of a variety of criteria, including the information source, attack detection methodologies, and performance metrics. Finally, current study gaps in the literature have been highlighted, and this will serve as a benchmark for future IIoT threat detection researchers.**

Keywords—Industrial IoT; Cybersecurity; Attack detection;Feature Extraction; Classification

I. INTRODUCTION

The "Mechanisation of Anything and everything" is indeed the product of a contemporary industrialization that offers profound transformation and human wellbeing. It connects digital gadgets, data mining, and real-world application administration via networked computers [1]. This revolution's possibility allows everyone to have accessibility to billions of dollars worth of knowledge and information, which offers new possibilities [2] [3].

The Four Key Components of Industrial IoT Architecture: 1. Intelligent Edge Gateway": It is a software application which can gather, combine, as well as sanities lightweight data flowing and therefore is tightly connected with sensor network. It enables aggregated and relevant data to be uploaded to the IoT network. It acts as a liaison between the equipment and also the larger cloud IoT network. 2. IoT Cloud": The basic IoT infrastructure that handles huge amounts of data using data processing, machine learning, and artificial intelligence approaches. " Device control, stream analytics, event management, a rules engine, alerts, and updates are all" accessible processing capabilities. " Big data analytics, as well as authorization, virtualization, end-to-end encryption, SDKs, and application APIs", are all available. 3."Business Incorporation and Platform": It is a backbone architecture that combines a multitude of IT strategies in needed to guarantee that computers data is gathered and analyzed throughout the whole operating cycle. ERP, Quality Management Systems, "Planning and Scheduling" Systems, and other systems are examples of such systems. 4. Determined by the shape of the output received, data analysis may be split into three divisions. Descriptive, predictive, and prescriptive analysis are the three forms of analysis. Figure 1 shows the four (4) levels of IIoT architecture, which are devices, intelligent gateways, IoT clouds, and business applications and connectors [4][5][6].

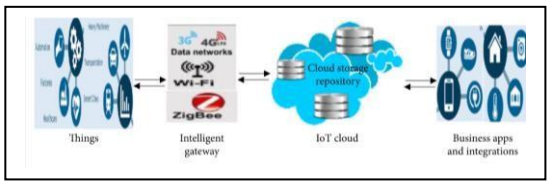


Fig. 1. Levels of IIoT architecture

The Industrial Internet of Things (IIoT) is known for generating large amounts of information from numerous sensors. Healthcare, retail, automotive, as well as transportation have all been impacted by such applications. The IIoT has the potential to boost effectiveness, productivity, and operational excellence inside a multitude of areas. Many businesses understand how and where IIoT innovations and solutions may emerge in organizational transformations, new and enhanced commodities / services, and whole innovative business models. By merging technical breakthroughs, sensors, programmers, and applications on the IIoT, machine learning and deep learning algorithms can improve dependability, performance, and customers ' satisfaction. In IIoTs, "machine-to-machine (M2M) and machine-to-person (M2P)" network connection is established utilizing the TCP/IP interface and several IIoT protocols [9] [10]. The number of flaws and defects that may be exploited using a variety of sophisticated methods of attack has expanded dramatically as the number of IIoTs has proliferated. Assailants try to exploit such systems in order to steal sensitive information, commit financial resources, and alter device resources [11]. As the number of IIoT devices and installations begins to proliferate, safeguarding essential services and infrastructure has become

a much more pressing issue for any business [13]. Malware exploiting zero-day vulnerabilities is one of most common threats in IIoT networks. Using tactics such as "Progressive Determined Risk (PDR), Denial-of-Service (DoS), and Decentralized DoS (DDoS)", the criminals infect susceptible machines in order to track and influence their behaviors . In response to the challenges represented by such intrusive frameworks, academics have indeed been encouraged to develop novel IDSs. Several intrusion detection systems (IDS) have already been created as well as enhanced in the past, but still they remain prone to a range of attacks. The potential of IDS to observe and foresee hostile conduct and unknown attacks has sparked a surge in interest in intrusion detection system research. Current machine learning-based irregularity detection algorithms, on the other hand, have such a substantial false alarm rate [8]. According to recent studies, feature extraction is now at the heart of much more reliable IDS [10 [11][12]. Classifier feature vectors, in particular, are large, but not all of them relate to the categories to be classified, necessitating the adoption of a feature selection approach. The machine learning based techniques have been utilized in most of the existing techniques for instruction detection in IIoT. But, still there aren't accurate outcomes. Therefore, the recent researchers have shifted towards the deep learning techniques, as they are good in enhancing the detection accuracies. Moreover, these classifiers are trained with the features acquired from the collected data. Therefore, it's essential to select the most suitable features.

Overall goal: To learn about just the present threats and remedies for security-related challenges in the Industrial IoT, as well as to identify studies that need to be done in this area. We shall conduct a comprehensive review of related literature to attain our aim.

The major contribution of this research work is:

- ✓ To review the recent works undergone in IIoT attack detection
- ✓ To summarize the concepts of the collected research papers comprehensively
- ✓ To analyze the collected research paper in diverse aspects like data source, pre-processing technique, feature extraction technique, feature selection technique and classifiers used.
- ✓ To discuss the advantage and drawback of each of the collected research papers
- ✓ To address the recent gaps in IIoT attack detection.

The left over section of this paper is arranged as: Section II discusses the recent work on IIoT attack detection and Section III manifests the information regarding the analysis on the collected IIoT attack detection papers. The research gaps and challenges identified are manifested in Section IV. This paper is concluded in Section V with future scope.

A. Motivation

Antivirus software and intermediary boxes, such as "Intrusion Detection Systems (IDS)" have been used in security systems. IDSs (intrusion detection systems) are security monitoring tools. They monitor the system for fraudulent attacks and examine network activity. In particular, when a harmful event is discovered, IDSs alert the network administrator. IDSs frequently employ abuse, anomalous, other hybrid detection techniques. Unauthorized assaults are identified by knowledge rules with misuse identification. A hypothesis is used to correlate assailant activity to normal behavior in abnormality identification. Abuse and anomaly mechanisms are combined in methodologies. For anomaly detection in the IoT, a variety of machine learning approaches have been developed. Machine learning-based and deep learning methods have been shown to be useful in detecting aberrant network traffic flow occurrences.

II. LITEARTURE REVIEW

A. Related Works

In 2021, Basset et al. [1] have developed a "forensics-based DL model (called Deep-IFS)" with the intention of identifying the IIoT traffic intrusions. The "local gated recurrent unit (LocalGRU)" has been utilized to grab knowledge on the local representations. Moreover, they have learnt the global representation with the MHA layer. The Big IIoT traffic data handling has been a major challenge, and this was addressed with the proposed Deep-IFS. The research has been dedicated for Bot-IIoT attack detection, and the results acquired have been showed that the projected model has been highly robust over the existing models.

In 2019, Zolanvari et al. [2] have cosnucted a cyber-vulnerability assessment and have discussed about the vulnerabilities that has been solved by the machine learning models. Following that, they have developed a new intrusion detection system (IDS) that has been applicable for detecting the, "backdoor, command injection, and Structured Query Language (SQL) injection attacks".

In 2020, Rubio et al. [3] have investigated the feasibility of machine learning models inside the IIoT for detecting attacks by examining their implementation across various IIoT infrastructures and establishing a standard platform for data collecting that takes into account the processing restrictions. As a consequence, a useful understanding has emerged, demonstrating the viability of such a method whenever deployed to future IIoT architectures.

In 2020, Li et al. [4] have developed a new "bidirectional long and short-term memory network with multi-feature layer (B-MLSTM)" for IIoT attack detection. In the models learning phase, sequence and stage feature layers have been implemented, which could also understand the relevant attacker frequency using past information, allowing the system to efficiently identify assaults involving different session. The classification method was therefore updated using a double-layer reverse device. The retraining time has been dynamically determined to meet the new attacking frequency by gathering information from testing data and doing unsupervised approach with statistical information.

In 2021, Kasongo et al. [5] have projected a new Intrusion Detection Systems (IDSs) for Industrial IoT (IIoT) attack detection. The feature selection has been carried out with the

Genetic Algorithm (GA). The intrusion detection has been accomplished with the machine learning models like "Random Forest (RF), Linear Regression (LR), Naïve Bayes (NB), Decision Tree (DT), Extra-Trees (ET), and Extreme Gradient Boosting (XGB)". The projected model ahs been validated with the data collected from "UNSW-NB15", the experimental outcomes have exhibited the supremacy of the projected model in terms of test accuracy (TAC) and Area Under the Curve (AUC).

In 2020, He et al. [6] have developed "BoSMoS-blockchain-based software status monitoring system" for IIoT malicious behaviors detection. The software states were monitored and recorded by BoSMoS. In addition, the software state were recorded towards guaranteeing the integrity information by the blockchain as a distributed ledger. Different consensus methods can be used on the BoSMoS blockchain network. They also assess BoSMoS's results in terms with exception reaction times, intrusion security, and adaptability. The experimental findings show that BoSMoS is both practicable as well as reliable.

In 2021, Chen et al. [7] have developed Fiden as a technique that fingerprints heterogeneous IIoT devices without taking regularity into consideration. This technique for analyzing from such a time - series data of data transfer and afterwards clusters the features to determine a device's fingerprint. A practical research model on the communications environment of the automotive sector has been used to show Fiden's applicability. The findings indicate that the suggested approach aids in the detection of device-mounted assaults.

In 2022, Lu et al. [8] have proposed Cognitive Memory-guided AutoEncoder (CMAE), an unique technique based on deep neural networks for tackling the intrusion detection challenge. The memory module has been used by the CMAE approach to strengthen the capable of storing regular feature patterns while retaining the autoencoder's benefits. As a result, this could withstand specimens that are unbalanced. Furthermore, detecting cyberattacks using the reconstruction error as an assessment parameter efficiently identifies unknown threats. They proposed "feature reconstruction loss and feature sparsely loss" to restrict the suggested memory module, increasing the divided into various types of storage elements as well as the capacity of depiction for normal data, in order to achieve a good intrusion detection performance.

In 2022, Hawawreh et al. [9] have developed a holistic approach called X-IIoTID for IIoT intrusion detection. For fitting the "heterogeneity and interoperability" of IIoT systems, the connectivity- and device-agnostic incursion dataset has been employed. It has multi-view characteristics including such "network traffic, host resources, logs, and warnings", and provides an assault categorization. To prove its originality, X-IIoTID was indeed analyzed utilizing common machine and deep learning methods and contrasted to 18 intrusion datasets.

In 2020, Latif et al. [10] have developed a new "deep random neural (DRaNN)-based technique" for IIoT intrusion detection. The UNSW-NB15 IIoT security dataset has been used to test the projected model. The suggested model effectively categorized 9 multiple kinds of assaults with a low false-positive rate and a good precision of 99.54 percent, according to empirical observations. With a detection rate of 99.41%, the suggested model outperformed the competition.

In 2020, Latif et al. [11] have projected a new "lightweight Dense Random Neural Network (DnRaNN)" for IIoT intrusion detection. Because of its inherent better generalization powers as well as decentralized nature, the suggested approach is generally preferred for deployment in resource-constrained IIoT networks. The suggested study's conclusions give suggestions as well as perspectives into binary and multiclass settings.

In 2019, Aydogan et al. [12] have utilized the IPv6 Routing Protocol for "Low-Power and Lossy Networks (RPL)" for Industrial IoT attack detection. Because the hackers use the RLP protocol's unlawful parental selection algorithm, cyberattacks targeting RPL have indeed been demonstrated to be conceivable. They presented a methodology and architecture for detecting intrusions into the IIoT in this research. Their findings show that the site successfully identifies "routing attacks in RPL-based Industrial IoT networks".

In 2021, Alcazar et al. [13] Have conducted a thorough analysis of Differential Privacy (DP) strategies used throughout the training of an Industrial IoT-enabled IDS using Federated Learning (FL) . They investigated the efficiency gained using alternative confidentiality constraints and consolidation methods, notably FedAvg as well as the recently proposed Fed+, using non-iid data from the latest ToN IoT dataset. Also when noise has been introduced in the federated training process, Fed+ tends to show similar outcomes in their context, as per the assessment.

In 2021, Nayak et al. [14] have developed a reliable DL-based routing attack detection technique for IIoT attack detection. In Routing Protocol in Low-Power and Lossy Networks, they investigated adversarial training of the system towards identifying planned assaults. The above assists in the development of a dependable learning approach. For attack detection and prevention activities, a "two-stage combination of GAN and SVM" models called the "Generative Adversarial Network-Classifier (GAN-C)" has been created. They compared GAN-C to a solo "Support Vector Machine (SVM)" classifier to see how much better it performs. The suggested technique uses a parallel learning and detection model to facilitate DL on IIoT devices with limited computing resources.

In 2021, Fang et al. [15] have developed a very effective weak EMI assault detection approach for IIoT attack detection. The projected model has been based upon the fingerprint of an intelligent device. The fingerprint has been retrieved using the Kalman technique and the Linear Time-Invariant (LTI) model. Next, a fusion model was developed to detect if the equipment is assaulted by weak EMI based on the retrieved fingerprint. To increase detection performance, the fusion model integrates the "Feature Extraction Unit (FEU)" with the "Long Short-Term Memory (LSTM)". Finally, an edge computing framework has been developed to improve the method's effectiveness.

The advantages and drawbacks of the existing models is manifested in Table I.

TABLE I. ADVANTAGES AND DRAWBACKS OF COLLECTED RESEARCH WORKS

Author[Citation]	Advantages	Drawbacks
Basset et al. [1]	✓ Deep-IFS reduces the danger of gradient vanishing by parallelizing the learning calculation (GPU execution), which is impossible with regular RNNs. ✓ Deep-IFS makes data transmission across fog nodes easier and reduces overheads, resulting in a valuable decision ✓ Support framework that helps individuals and IIoT service providers share their data in a trustworthy and safe manner.	The Deep-IFS is taught in a responsibility for protecting, preventing it from learning from unlabeled traffic. Second, the suggested framework fails to address how data privacy would be maintained, which is a critical feature of sensitive industrial applications. Large amounts of IIoT traffic may reduce Deep-efficiency IFS's in modifying incoming traffic with no misses.
Zolanvari et al. [2]	✓ consumes lower time	less accuracy higher false negative
Rubio et al. [3]	✓ Addressing all of the project's structural and computational limitations.	reduced security
Li et al. [4]	✓ lower Mean square error (MSE)	Lower misclassification accuracy Higher computational complexity
Kasongo et al. [5]	✓ higher classification accuracy	higher prediction time
He et al. [6]	✓ Can be used in a large-scale IIoT environment. ✓ Response time is reduced. ✓ It capable of successfully safeguarding software confidentiality and achieving consistency	Higher time consumption
Chen et al. [7]	✓ Doesn't have to take signal transmission periodicity into account. ✓ When compared to CIDS, Fiden can determine the attacker's ECU.	lower security
Lu et al. [8]	✓ higher F1-score (F1) and accuracy, Recall (R)	Higher computational complexity. requires huge time
Hawawreh et al. [9]	✓ robust security	not applicable for huge database
Latif et al. [10]	✓ Higher accuracy and efficiency	Does not encompass cyber physical assaults on PLCs that have a direct impact on their physical characteristics. susceptible to over fitting
Latif et al. [11]	✓ designed for a multiclass problem	not too effective for the diverse nature of intrusions
Aydogan et al. [12]	✓ An intrusion detection system that is appropriate for RPL-based IIoT. ✓ It is capable of detecting intrusions with "excellent accuracy (high true positive and low false positive rates)".	Cannot diminish the attack's effects. increased computing and communication costs
Alcazar et al. [13]	✓ better accuracy	There are no privacy-preserving measures included. continues to have privacy concerns with the sharing of gradients/weights within every learning cycle
Nayak et al. [14]	✓ Be competent to efficiently do deep learning. ✓ This skill seems critical for identifying anomalous incursion features in network activity and distinguishing anomalous behavior from general traffic.	It is appropriate for smaller datasets. There aren't any trade-offs between detection settings.
Fang et al. [15]	✓ ensuring the security of IIoT systems	higher computational complexity

III. ANALYSIS ON THE COLLECTED RESEARCH WORKS

A. *Analysis on Database*

The different databases used in each of the research projects are shown in Fig.2. The dataset Bot-IoT was used in [1], UNSW-NB15 was used in [2], KDD CUP 99 was used in [7], [10], and [15], and KDD CUP 99 was used in [2], [7], [10], and [15].

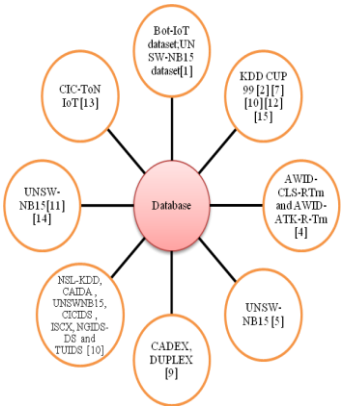


Fig. 2. Analysis on database collected in each resarch works

B. *Pre-processing Techniques*

Fig.3 shows the various pre-processing strategies employed in the gathered research papers. In most research studies, data normalization has been the recommended method. [4] [7], [9], [10], [11], [14].

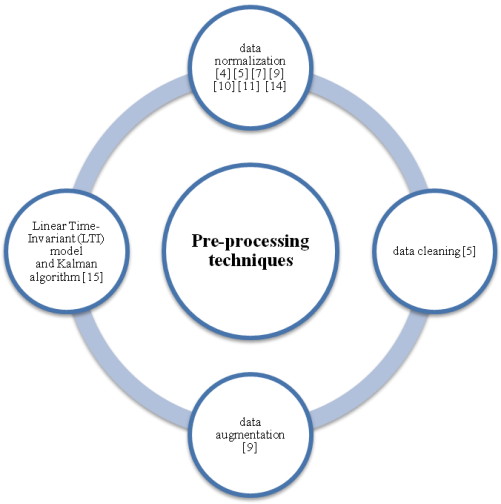


Fig. 3. Analysis on pre-processing techniques utilized in each resarch works

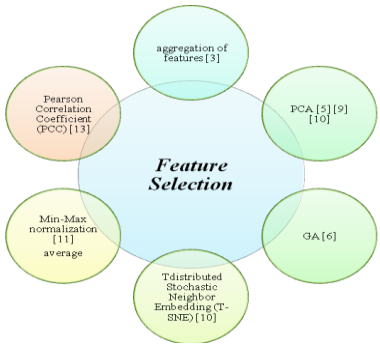
C. *Feature extraction and Feature selection*

This section discusses the various feature extraction strategies as well as the characteristics that were used to train the classifier for assault detection. A clear review of the feature extraction techniques is shown in Table II. In [2], the mean flow, source port, destination port, destination packets, total packets, source bytes, destination bytes, total bytes, source load, destination load, total load, source rate, destination rate, total rate, source loss, destination loss, total loss, total percent loss, source jitter, destination jitter, source interpacket,destination interpacket features have been extracted. In [11] and [12], the flow based feature has been extracted. Furthermore, the feature selection approaches that were used in the gathered study article are addressed.

TABLE II. ANALYSIS ON FEATURE EXTRACTION TECHNIQUES

Author[Citation]	feature extraction
Basset et al. [1]	G1 and G2 feature: attention mechanism
Zolanvari et al. [2]	mean flow, source port, destination port, destination packets, total packets, source bytes, destination bytes, total bytes, source load, destination load, total load, source rate, destination rate, total rate, source loss, destination loss, total loss, total percent loss, source jitter, destination jitter, source interpacket, destination interpacket
Rubio et al. [3]	"Number of connections established and devices accessed Traffic load (total number of packets exchanged) Type of communication protocols used Delay experienced in every communication channel Ratio of lost/corrupted packets Frequency and type of commands issued Precise data values transmitted by sensors"
Li et al. [4]	network traffic data
Kasongo et al. [5]	statistical features; G!, G2, G3, G4, G5, G6 and G7
Chen et al. [7]	"features of signal transmission, timestamp and accumulated clock offset"
Hawawreh et al. [9]	connectivity-agnostic features
Latif et al. [10]	"edge gateway's resources, such as its CPU and memory loads, I/O activities and the system's load, process and context switch"
Latif et al. [11]	flow based feature
Aydogan et al. [12]	flow based feature
Alcazar et al. [13]	flow ID;destination IP addresses
Nayak et al. [14]	TensorFlow data structure
Fang et al. [15]	Long Short-Term Memory (LSTM)

The feature selection techniques used in the existing research papers is shown in Fig.4.



D. *Attack detector*

The different "machine learning and deep learning techniques" that has been utilized for attack detection is discussed in this section. The analysis on the classifiers used

in the collected research papers is manifested in Table III. In [5], the classifiers like "RF, Linear Regression (LR), Naïve Bayes (NB), Decision Tree (DT), Extra-Trees (ET), and Extreme Gradient Boosting (XGB)" has been used for attack detection. In most of the works, the machine leraning techniques have been used. But, still they lack in detection accuracy.

TABLE III. ANALYSIS ON CLASSIFERS USED IN COLLECTED RESERCH WORKS

Author[Citation]	classification
Basset et al. [1]	Deep-IFS Approach
Zolanvari et al. [2]	SCADA-RF, decision tree, KNN, Logistic regression, SVM, SNN and Navies Bayes
Rubio et al. [3]	Opinion Dynamics algorithm
Li et al. [4]	"bidirectional long and short-term memory network with multi-feature layer (BMLSTM)"
Kasongo et al. [5]	"RF, Linear Regression (LR), Naïve Bayes (NB), Decision Tree (DT), Extra-Trees (ET), and Extreme Gradient Boosting (XGB) "
He et al. [6]	BoSMoS
Chen et al. [7]	Fiden
Lu et al. [8]	DDoS
Hawawreh et al. [9]	X-IIoTID
Latif et al. [10]	DT, NB, KNN, SVM,DNN, LR and GRU
Latif et al. [11]	deep random neural (DRaNN)
Aydogan et al. [12]	RPL protocol genetic programming (GP)
Alcazar et al. [13]	generative adversarial networks (GANs)
Nayak et al. [14]	"Bi-directional Long Short-Term Memory Recurrent Neural Network (BLSTM RNN)"
Fang et al. [15]	CNN

E. Type of Attack

This section discusses the types of attacks that have been examined in each of the research studies. The type of attacks focused in literature works is manifested in Table IV. The DDoS attacks have been discussed in most of the research works [4] [6] [10] [11] [13], respectively.

TABLE IV. TYPES OF ATTACKS FOCUSED IN LITERATURE WORKS

Author[Citation]	attack type
Basset et al. [1]	BOTNET
Zolanvari et al. [2]	backdoor, comman injection and SQL injection attack
Rubio et al. [3]	Advanced Persistent Threats (APTs)
Li et al. [4]	DDoS attack
Kasongo et al. [5]	"Fuzzers, Analysis, Exploits, Worms, Shellcode, DoS, Generic, Reconnaissance, and Backdoor"
He et al. [6]	Distributed Denial of Service (DDoS)
Chen et al. [7]	DoS Attack;Fuzzy Attack; Impersonation Attack;
Lu et al. [8]	Cognitive Memory-guided AutoEncoder (CMAE)
Hawawreh et al. [9]	"Reconnaissance:;Weaponization;Exploitation;Lateral Movement; Command and Control (C&C);Exfiltration;Tampering; Crypto-Ransomware"
Latif et al. [10]	Distributed Denial of Service (DDoS)
Latif et al. [11]	DDoS attack
Aydogan et al. [12]	Hello flood attack; Version number attack;
Alcazar et al. [13]	"DoS, DDoS, Backdoor, Injection, MITM, Scanning, Password, and XSS"
Nayak et al. [14]	Backdoor, Denial-of-Service;Worms;Reconnaissance
Fang et al. [15]	weak electromagnetic interference attacks

F. Performance Analysis

The performance recorded by each of the research work is manifested in Table V. In [6], the Total throughput=28,265,152 and in [15], the Average running time (s) =0.013. In [11], the best value detection rate of 99.41%. Has been recorded, and this improvement is owing towards the usage of the deep learning classifier- deep random neural (DRaNN).

TABLE V. PERFORMANCE RECORDED BY THE LITERATURE WORKS

Author[Citation]	performance
Basset et al. [1]	Accuracy: 99.77; Precision: 99.99; Recall: 99.77; F1-measure: 99.88
Zolanvari et al. [2]	Accuracy recorded: RF=99.9, decision tree=99.98, KNN=99.98, Logistic regression=99.90,SVM=99.64, SNN=99.64 and Navies Bayes=97.48; Average data rate=419 kBits/sec
Rubio et al. [3]	Accuracy=97.58%,detection rate=83.79%,FNR=6.02%
Li et al. [4]	Accuracy=97.58
Kasongo et al. [5]	test accuracy (TAC) of 87.61% and an Area Under the Curve (AUC) of 0.98
He et al. [6]	Total throughput=28,265,152 for Payload Size (bytes)=6K
Chen et al. [7]	average accuracy of Fiden-timestamp is 0•936507937, Fiden-offset is 0•936507936, and CIDS is 0•813492064
Hawawreh et al. [9]	Detection rate (%) =99.97
Latif et al. [10]	For 9 attacks, it achieves 35.52%, 55.87%, 75.20% and 47.96% for accuracy, precision, recall and the F1-score, respectively
Latif et al. [11]	great accuracy of 99.54%; Detection rate of 99.41%.
Aydogan et al. [12]	accuracy higher than 90%
Nayak et al. [14]	Accuracy=0.9571%; Recall=0.96%;f1 - score=0.98%; Miscalculation rat=0.041%; Detection Time (sec)=2.19%
Fang et al. [15]	Average running time (s)=0.013 for fingerprint length=45

IV. RESEARCH GAPS AND CHALLENGES

Intrusion detection inside the era of the Industrial Internet of Things remains a challenge, according to this investigation. The emphasis switches from connections to information as the World Wide Web matures into the Internet of Things. As a result, the focus of this research was on the most recent research in intrusion detection and intelligent algorithms used to IoT to keep data safe. The works reviewed in this study largely focused on the concerns as well as various efforts made by the scientific community as well as industry to design efficient security procedures that provide acceptable protection while consuming very little energy. Although such strategies aim to improve intrusion detection identification rates, it is believed that the false positive rate and accuracy remains an problem that has to be addressed in all the upcoming research. While some strategies [2][8] can minimize the percentage of false positives, they can increase training time and categorization. Some techniques, on the other hand, conduct the inverse process, stabilizing the false positive rate at the cost of a considerable computational load during training and testing. This is especially important in the case of intrusion detection, when real-time detection is important. Machine learning techniques have been shown to deliver the best intrusion detection accuracy. Machine learning algorithms produce better outcomes than other techniques even though they can be implemented to a variety of databases and therefore can examine real-time data. Using several machine learning algorithms, including such logistic regression, NB, DT, KNN, and RF, a trust model for machine-to-machine communication was created in a prior work [4]. To determine the optimal strategy, a comparison research was conducted [5]. That research looked into a variety of approaches, including Nave Bayes, SVMs, and decision trees. This technique allows credible data on unusual activities but may also be used to determine the source of the intrusion or the primary problem. Such issues are often recognized derived from data patterns that take a lot of time for human experts. Large data sets have been studied in this research, which is labor-intensive and time-consuming using traditional methods. CNN, CNN-LSTM, CNN-RNN, and CNN-GRU are instance of deep learning algorithms that have already been employed to detect intrusions. These methods have shown it to be more effective; nevertheless, because to the complicated design, training incurs a large computational cost. An Deep-IFS model was built to improve accuracy [1]. When the suggested LocalGRU model was compared to an SVM, the findings showed that the proposed model was more accurate. The hybrid technique of GAN with an SVM has been applied against network intrusion [14]. This method proved to be far more accurate than using an SVM alone. The disadvantage of this methodology is that it generates more false positives than other approaches, such as BPN.

V. CONCLUSION AND FUTURE SCOPE

This research has undergone a comprehensive conceptual evaluation with research publications on recent IIoT threat detection efforts. The database utilized, pre-processing approach, feature selection, classification (machine learning/deep learning), and performance reported have all been examined in these publications. The research gaps revealed in the existing models have also been discussed. According to the results of the survey, it was discovered that the most acceptable characteristics for training the model must be extracted. Furthermore, in order to decrease the computational complexity that has been encountered in most studies, it is necessary to choose the best characteristics from the retrieved ones. The classifier then recognizes the final conclusion regarding the existence or absence of assaults in the network. Since machine learning models have been shown to be less accurate in the literature, deep learning models with optimization algorithms are recommended as a possible technique for improving detection accuracy. Furthermore, because very few efforts have been focused on attack mitigation, a new attack mitigation model is required.

REFERENCES

[1] M. Abdel-Basset, V. Chang, H. Hawash, R. K. Chakrabortty and M. Ryan, "Deep-IFS: Intrusion Detection Approach for Industrial Internet of Things Traffic in Fog Environment," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7704-7715, Nov. 2021. doi: 10.1109/TII.2020.3025755

[2] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan and R. Jain, "Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things," in *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822-6834, Aug. 2019. doi: 10.1109/JIOT.2019.2912022

[3] J. E. Rubio, R. Roman and J. Lopez, "Integration of a Threat Traceability Solution in the Industrial Internet of Things," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6575-6583, Oct. 2020. doi: 10.1109/TII.2020.2976747

[4] X. Li, M. Xu, P. Vijayakumar, N. Kumar and X. Liu, "Detection of Low-Frequency and Multi-Stage Attacks in Industrial Internet of Things," in *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8820-8831, Aug. 2020. doi: 10.1109/TVT.2020.2995133

[5] S. M. Kasongo, "An Advanced Intrusion Detection System for IIoT Based on GA and Tree Based Algorithms," in *IEEE Access*, vol. 9, pp. 113199-113212, 2021. doi: 10.1109/ACCESS.2021.3104113

[6] S. He, W. Ren, T. Zhu and K. R. Choo, "BoSMoS: A Blockchain-Based Status Monitoring System for Defending Against Unauthorized Software Updating in Industrial Internet of Things," in *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 948-959, Feb. 2020. doi: 10.1109/JIOT.2019.2947339

[7] Y. Chen, W. Hu, M. Alam and T. Wu, "Fiden: Intelligent Fingerprint Learning for Attacker Identification in the Industrial Internet of Things," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 882-890, Feb. 2021. doi: 10.1109/TII.2019.2962759

[8] H. Lu, T. Wang, X. Xu and T. Wang, "Cognitive Memory-Guided AutoEncoder for Effective Intrusion Detection in Internet of Things," in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3358-3366, May 2022. doi: 10.1109/TII.2021.3102637

[9] M. Al-Hawawreh, E. Sitnikova and N. Aboutorab, "X-IIoTID: A Connectivity-and Device-agnostic Intrusion Dataset for Industrial Internet of Things," in *IEEE Internet of Things Journal*. doi: 10.1109/JIOT.2021.3102056

[10] S. Latif, Z. Idrees, Z. Zou and J. Ahmad, "DRaNN: A Deep Random Neural Network Model for Intrusion Detection in Industrial IoT," *2020 International Conference on UK-China Emerging Technologies (UCET)*, Glasgow, UK, 2020, pp. 1-4. doi: 10.1109/UCET51115.2020.9205361

[11] S. Latif *et al*., "Intrusion Detection Framework for the Internet of Things using a Dense Random Neural Network," in *IEEE Transactions on Industrial Informatics*. doi: 10.1109/TII.2021.3130248

[12] E. Aydogan, S. Yilmaz, S. Sen, I. Butun, S. Forsström and M. Gidlund, "A Central Intrusion Detection System for RPL-Based Industrial Internet of Things," *2019 15th IEEE International Workshop on Factory Communication Systems (WFCS)*, Sundsvall, Sweden, 2019, pp. 1-5. doi: 10.1109/WFCS.2019.8758024

[13] P. Ruzafa-Alcazar *et al*., "Intrusion Detection based on Privacy-preserving Federated Learning for the Industrial IoT," in *IEEE Transactions on Industrial Informatics*. doi: 10.1109/TII.2021.3126728

[14] Sharmistha Nayak, Nurzaman Ahmed, Sudip Misra,"Deep Learning-Based Reliable Routing Attack Detection Mechanism for Industrial Internet of Things", Ad Hoc Networks, 2021

[15] Kai Fang,Tingting Wang,Jianqing Li,"Detection of weak electromagnetic interference attacks based on fingerprint in IIoT systems",Future Generation Computer Systems,2021