Bachelor's thesis

Information and Communications Technology

2023

Mohammed Al-Humairi

Developing an Efficient Attack Detection Model for an Industrial Control System using CNN-based Approaches

– attack detection using PS-CNN

TURKU AMK
TURKU UNIVERSITY OF APPLIED SCIENCES

Bachelor's Thesis | Abstract

Turku University of Applied Sciences

Information and Communications Technology

2023 | 37

Mohammed Al-Humairi

Developing an Efficient Attack Detection Model for an Industrial Control System using CNN-based Approaches
– attack detection using PS-CNN

For uninterrupted operations, Industrial Control Systems (ICS) critical infrastructures such as thermal plants, water treatment plants, nuclear plants, oil refineries, and gas pipelines rely on ICS components. These systems include Supervisory Control and Data Acquisition (SCADA) systems that gather and monitor real-time data. ICS has evolved from proprietary to open architectures which are highly dependent on computer networks and the internet. This transformation exposes these systems to novel threats and cyber-attacks. As security is a major concern, this thesis aimed to design an intelligent attack detection model for cyber-attack detection in an ICS with high detection accuracy rate and fewer false positives. Hence in this thesis, an effective Attack Detection model namely PCA Select KBest CNN (PS-CNN) designed for detecting attacks in the ICS with better accuracy in detection and lesser false positives was developed. Further, the informative features were identified from the selected UNSW-NB15 dataset using the Principle component Analysis for feature extraction and the Select KBest approach for the feature selection process. Finally, malicious attacks in ICS with higher accuracy in attack detection were classified using Convolutional Neural Network- (CNN). The results and comparative analysis show the significance of the proposed attack detection model over existing state-of-the-art techniques.

Keywords:

Industrial Control Systems, Cyber attacks, Convolutional Neural Network

Contents

Equations

Figures

Tables

List of abbreviations

ACCS	Australian Centre for Cyber Security
AML	Adversarial Machine Learning
CNN	Convolutional Neural Network
IoT	Internet of Things
ICS	Industrial Control System
IACS	Industrial Automation and Control Systems
PCA	Principal Component Analysis
PSOM	Probabilistic Self-Organizing Maps
PS-CNN	PCA Select KBest CNN
SCADA	Supervisory Control and Data Acquisition

1 Introduction

1.1 Industrial Control System

Cyber security is considered as Industrial automation and control systems which have control over Supervisory Control and Data Acquisition (SCADA) network and its support systems (Fig 1). It differs from the field level to the control level and then from the plant to the security level. There is an increasing concern regarding these kinds of cyber-attack which result in the integration of IACS and Convolutional Neural Networks. Thus. various types of proposed frameworks are categorized into feature extraction model and isolation forest model for the detection. The basic structure is shown in the following Fig.1. Industrial automation and control systems (IACS) are profusely employing supervisory control and data acquisition (SCADA) networks. However, their integration into IACS is vulnerable to various cyber-attacks (Elnour, Meskin et al. 2020).

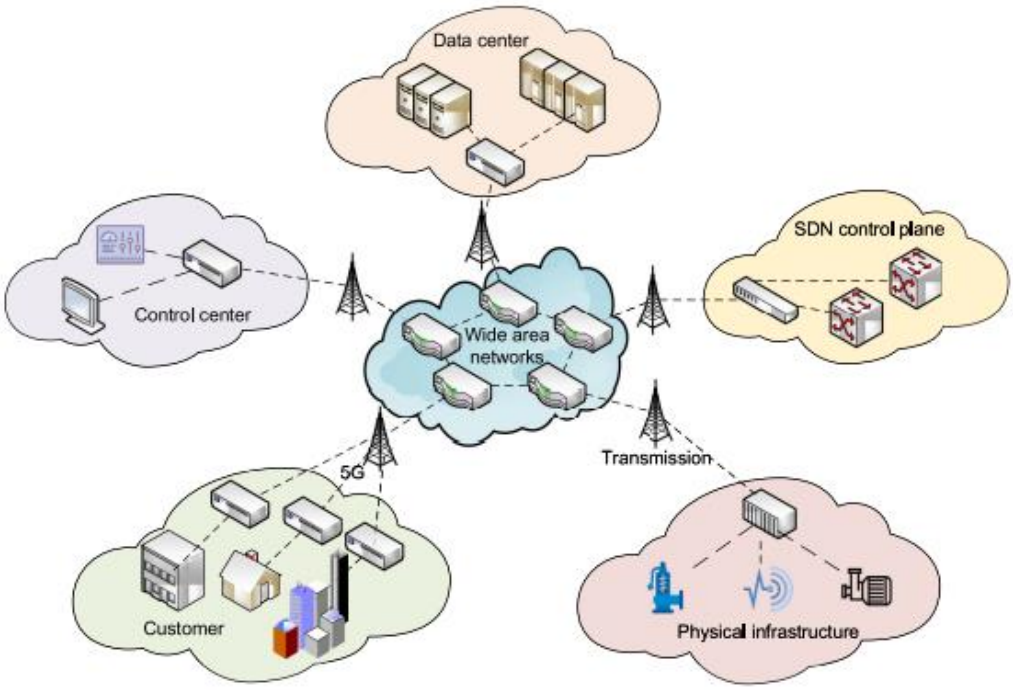

Figure 1. Basic Architecture of Industrial Control Systems

Source : (Li, Wu et al. 2020)

It has also been reported (Zyad, Taha et al. 2019) that many threats can harm environmental security thereby insisting that network invasions with increased frequency consequently promoting easy access to automated attack tools. Several kinds of methods have been employed which could highlight the variety of high to low dimensional space in which the Principal Component Analysis (PCA) is highly noted. This detection technique (anomaly and misuse based) is possibly divided into two categories where the first approach deals with network intrusions in which a set of predefined attack signatures is used and the latter deal with the familiarity mode of normal behaviour. The goal of this kind of model is to elaborate the feature space and distinguish between normal and abnormal patterns which are based on probabilistic self-organizing maps (PSOM).

Lu et al. (2021) state that the implementation of supervisory control and data acquisition leads to better performance of the ensemble learning scheme for single-method cyber-attack detention. This was evaluated using a gas pipeline system and a water storage system to improve the SCADA network with the comparison to other existing methods. All these methods have been infrastructured for promoting and detecting cyber security problems via IACS thus promoting the deep belief network as one of the best methods for identifying classification models. These methods are chosen to address the difficulties in cyber security using a surrogate-assisted model such as the Gaussian regression process or even Bayesian optimization. A neural network of Long Short Term memory could be used to further validate the sophisticated deep learning approach using different sets of data scales.

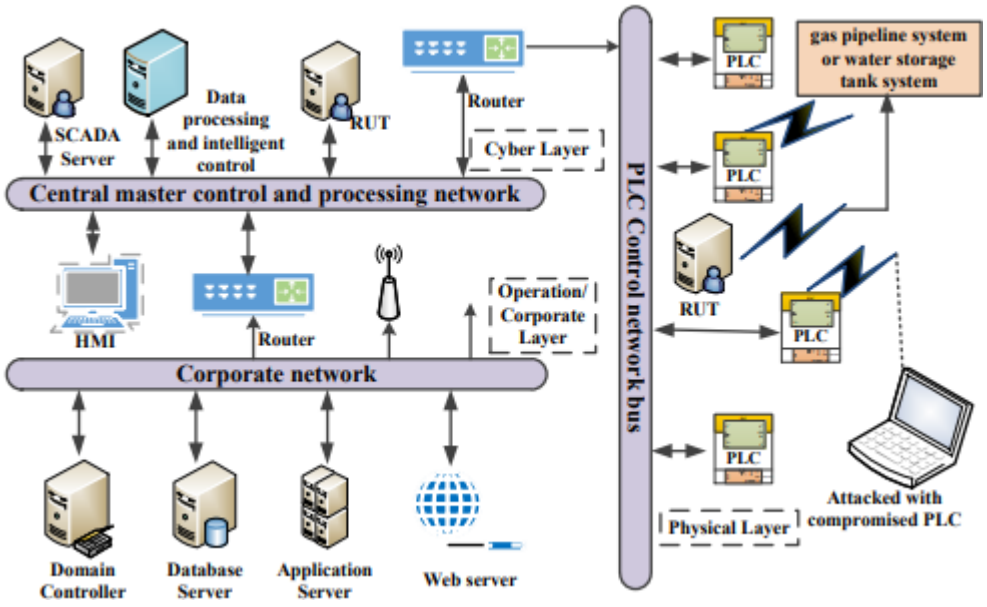

Figure 2. SCADA-based general structure using IACS

Source : (Lu, Zeng et al. 2021)

Abana et al. (2016) accounted for the reason for the increase in the mobile traffic data where demand for mobile applications and the increase of exponential increase is witnessed. Improvising the revolutionary methods of addressing the 5th generation procedures such as long-term evolution and LTE advanced systems increases the capability of more network data to sustain and make an efficient way for user experience. It was also found that D2D communication is another new approach that could address the increasing capacity demands where processing sufficient storage capabilities is highly possible. The achievement of this type of approach in addition to the non-uniform D2D scheme shows that this acquires low coverage of network data.

The internal or external operations of ICS may be affected by the following:

1. Changes in the system's operation or application configurations may produce unpredictable outcomes. This could be done to disguise malware or other kinds

of malicious activity. This could also have an impact on the output of the targeted attack.

2. Changes or adjustments in the ICS components and other controller devices can cause malfunction and loss of control over a process.

3. Misinformation may report to the function of the system. This event may result in the execution of unnecessary actions and changes to the programmable logic. This scenario sets a base to mask malicious activity, such as malicious code injection.

4. When the safety controls are tampered with, the lives of employees and external clients are put in danger.

1.2 Deep learning approaches for ICS attack detection

Dealing with different anomaly inhibition as industrial devices have limited resources and require legacy-compliant protocols, integrating contemporary security measures into existing ICS is difficult (i.e., they have to be backwards compatible to decades for old devices in the field, which do not support authentication or encryption). Adversarial machine learning (AML) is a crucial tool to investigate the resistance of machine-learning-based anomaly detectors to manipulations. Regarding the image, speech, and virus detection, several evasion attacks defence strategies have been presented. The limitations and breadth of the attacker change depending on the scenario. The attacker's objective in the instance of picture recognition could be to misclassify the sample, either on an intended target class or a random target class. The attacker's objective in malware detection is to misclassify a malware sample because the task is binary (malware vs. innocuous software). The restriction on the adversarial example is to maintain the behaviour of the malware, which means that the distortion made to the malware should not render it incapable of harm. A methodology of white and black box attacks results in the comparison like white box attack detection shows a zero value whereas the black box attack detection does not show any clearance in performance. (Erba, Taormina et al. 2019)

A deep learning -based intrusion model using CNN and a recurrent electric version and federated learning framework has been considered in many studies to detect a comprehensive intrusion model. In terms of an industrial context, the Internet of Things plays a dominant role in efficient grids, self-driving vehicles, and gas pipeline networks. The industrial control systems also include artificial intelligence, cloud computing, virtualization and software-defined networking. (Li, Wu et al. 2020).

The different detection methods are shown in Fig. 3, which had been featured by a specific positive rate and also pours out the failure effective rate on economics and human life. Data Acquisition System that connects the far and near networking protocols also cause an attack to threaten Industrial Control systems. Medical devices, transport and other areas of society assist in the use of industrial control systems which provides a reliable way for malware attack and their threats. The power grid in Ukraine and the Stuxnet malware attack in Iran on Saudi Oil Company are some recent examples of this kind of malware attack. SWAT, BATADAL and WADI datasets have been used as a supplementing tool that controls the sensory data in addition to the PCA (Principal Component Analysis) (Kravchik and Shabtai 2021).

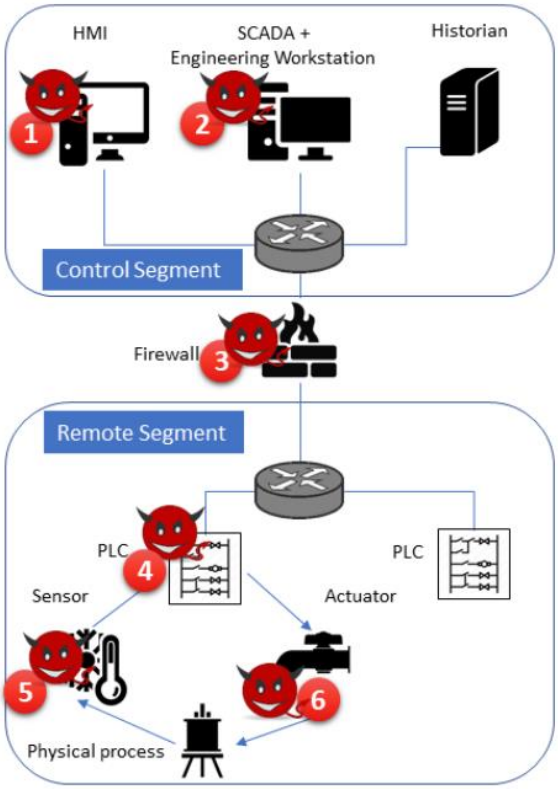

Figure 3. ICS architecture and specifications marking of malware attack

Source: (Kravchik, Biggio et al. 2021)

1.3 Problem statement

Over the previous decades, the developing technologies in ICS are continuously progressing which incorporating IT frameworks into physical components leads to security implications on water treatment plans, gas pipelines and smart grids. An ICS system usually focuses on availability rather than confidentiality and integrity. Normal and attack operation misclassification can restrict the ICS availability. Hence, the ICS principle provides an opportunity for attacks and enables the system to be at risk. In recent years, the number of cyber-attacks has increased and securing the technologies and system is not an easy task, as can be demonstrated from the existing research The existing research shows that ICS has been transformed into a hotspot for intrusion, attacks, vulnerabilities and threats. When considering the severity and criticality of cyberattacks against ICI,

earlier attack detection methods should be developed. Promising solutions are expected from research into ICS attack detection. Designing an effective ICS attack detection mechanism with higher accuracy in detection and a lesser false positive rate is said to be the best approach.

1.4 Objectives

The major goals of this thesis are,

- To design an effective Attack detection model namely (PS-CNN) PCA Select KBest CNN for detecting attacks in Industrial Control System with better accuracy in detection and lesser false positives.
- To identify the informative features from the selected UNSW-NB15 dataset using the Principle component Analysis for feature extraction and Select the KBest approach for the feature selection process.
- To identify malicious attacks in ICS with higher accuracy in attack detection, Convolutional Neural Network- CNN is implemented.

1.5 Thesis Organisation

Chapter 1 deals with the introduction to Industrial Control Systems (ICS), attack detection models using deep learning approaches, problem statement, and the objective of the study.

Chapter 2 discusses the state-of-the-art attack detection approaches for attack detection in ICS.

Chapter 3 deals with the design of an efficient feature extraction, selection approach, and classification process in ICS.

Chapter 4 deals with performance analysis for attack detection in ICS.

Chapter 5 deals with the overall conclusions and future scope of this research work.

2 Literature Review

Kravchik and Shabtai 2021 researched the different malware attack detection methods in which the proper protection of the ICS (Industrial Control System), becomes a source for the firms and the surrounding potentialities in the environment, also acts as an essential goal and a commodity for capital and survival of human beings. Depending upon the small convolutional neural networks such as ID convolutions and variable autoencoders, detection rates could be achieved either by analysing the minimal footprint, its training and detection rates or by the frequency domain analysis which is useful in anomaly and attack detection. It tested the proposed method on three well-known public datasets and achieve detection rates that are comparable to or superior to those reported in the literature, while showcasing minimal environmental impact, rapid training and detection, and generality. Additionally, the potency of PCA, which, with the right data preparation and feature selection, can yield excellent attack detection results in a variety of contexts is shown.

Kravchik et al. 2021 proposed on the neural network device which accounts for auto encoders collecting various operations focusing on industrial control systems. This helps the poisoning attacker in the cyber-attack which makes a white box attack scenario providing detailed information on the precisions of requirements and the wavelength. The three poisoning mitigation techniques had been analysed including the doorstep verge analysis, pattern distance extension, lowering the defined level, and employing dual sensors. To identify such poisoning factors, two components of theatrical methodologies had been involved in addressing the multivariate CNN regression-based tasks. Through this research, it was suggested that two different known attacks, interpolation and home-safe poisoning, show how powerful they are using both fictitious and actual ICS data. More than data sets such as test bed data (artificial and real), complex data of these algorithms could be used as the extended future study using principal component analysis and thereby increasing the depth of the extended algorithms.

Alladi et al. (2020) suggested that it is completely obvious that a hacker with insufficient funding would be unable to conduct focused attacks on industrial control systems. Separating this broad principle, it was widely discussed in the studies of ICS attacks over recent 20 years. The chosen attacks are the most significant in terms of the monetary destruction to finance, the potential for physical equipment damage, and the possibility of human losses. Moreover, the attack mechanism used in each of these attacks and suggested possible countermeasures were described. To gain a deeper understanding of the evolution of future cyber-security methods for ICSs, this case study had been progressed and analysed. More and more businesses that deal with hazardous chemicals, such as paper and pulp mills, weeds, petroleum products, etc., are getting outfitted with software solutions as technology's influence on our lives grows. This leads to a more precautionary system of ICs and software solutions which may put in difficult situations if not monitored properly.

Zhang et al. 2021 insists on the use of computer crime systems or cyber-physical system security - CPSs which would moderate as intricate systems that incorporate control, transmission, and computing technology. CPSs are used widely today in smart grids, smart manufacturing, smart cities, and intelligent transportation. The integration of industrial control systems with contemporary communication technologies would, however, inevitably expose CPSs to greater security risks, which could seriously impair system performance or even result in CPS destruction. It had been stated that recent developments are surviving in industrial cyber-physical system security (ICPS) probably through denial-of-service (DoS) attacks and the deception attack, two common types of attacks, and present recent findings regarding attack detection, estimation, and control of ICPSs could also be designed in future.

Even though numerous key findings on the security problem of ICPSs have been made from a variety of angles, the protection of ICPs is now facing several entirely new challenges due to the security environment’s complexity.

Li et al. 2020 viewed the importance of the threat landscape of industrial cyber-physical systems which had been significantly increased as a result of the rapid

convergence of current manufacturing infrastructures with intelligent networking and computing technologies (5G, software-networking, defined networking and artificial intelligence including CPSs). However, there are not enough high-quality attack instances; defending against cyber security against such massive, intricate and heterogeneous industrial CPSs has proven to be incredibly difficult. These cyber-physical security schemes focus on the two ways of improvement in its landscape namely analysing cyber threats based on opposing the industrial threats and intrusion detection model which is highly infusible attached with a collaborative framework based on deep federated learning methodologies. Research on the real industrial processes had been infused using strategies such as the Paillier cryptosystem based on the protocol of secure communication which is to safeguard the confidentiality of various strategies used in the production process of training. Extensive studies on a genuine industrial CPS dataset show the suggested Deep Fed scheme's great performance in identifying several kinds of cyber risks to industrial CPSs and their advantage of over cutting-edge techniques.

Lai et al. 2019 demonstrated the impact of more use of information technology with its limitations and drawbacks thereby insisting on the importance of cyber security attacks replacing with the technology of anomaly detection for the protection of industrial control systems. Considering this as an old-dated technology and identifying their shortcomings for the timely analysis and the anomalies location, a method of deep learning representation with CNN is used. The one-dimensional data is used to make the mapping process easier and acts as a good substitute for model processing. SCADA plays an efficient role in analysing the deep learning method more proficient and efficient in model evaluation. The application of deep learning methods through convolutional neural networks could be effective in analysing industrial anomaly attacks and detention. As Industrial Control Systems became an important factor in several countries, it is sure to detect security at the control level of any field work. As an improvement from the fieldwork process, it is important to enable SCADA anomaly detection and other deep learning framework methods to enhance the industrial process smoother and easier.

Lu et al. 2021 proposed supervisory strategies in the field of Industrial automation and control systems. This leads to the integration focusing on different kinds of cyber-attacks. Shaking premises in the background of different SCADA-based Industrial Automation results in the application of PEO-based algorithms leads to the variants of PEO and population extremal optimization (PEO)-based deep belief network detection method (PEO-DBN) based networks. Thus PEO-DBN-based networks could be enhanced by the ensemble learning scheme which proposes a method called En PEO-DBN where there are no hidden units available to frame better parameters. Thus SCADA network offers platforms like gas pipelines and water storage tank systems to evaluate the findings of the proposed detection methods for the prediction of evaluation analysis. PEO is used in PEO- DBN to automatically optimise DBN parameters rather than other manual adjustments. In En PEO-DBN, three PEO-DBNs with various features learn the mapping function between traffic characteristics and various cyber-attack kinds individually. The final detection results are then obtained using a majority-vote system. An assisted surrogate model such as the Gaussian process or process of Bayesian optimization is suggested for the future work of evaluation fitness in this research.

Noorizadeh et al. 2021 approached the implementation of the Tennessee Eastman process and the closed-loop controllers in the PLCs of Siemens. There is an equipment strategy called man-in-the-middle structure which is used to identify the hackers' corruption through false data cyber-attacks injection that has been supported by PLCs. Various detection algorithms could be detected through different findings and the evaluation by analysing the test bed performance of cyber-attack systems. As the development of cyber-physical industrial control systems in network infrastructure has grown recently, it is essential to tackle the cyber security issues posed by these systems to ensure their dependability and secure operation in the face of malevolent assaults. Creating a test bed to provide real-time data sets for crucial infrastructure that would be used for real-time validation is one way to do this. Algorithms for detecting security attacks are needed to maintain a structure for cyber security methodology. Cyber-attacks using the man-in-the-middle (MITM) technique are implemented directly

on the PROFINET routing protocols, allowing the malicious hacker to alter the sensor measurements supplied to the PLC.

The objective of this study (Al-Abassi, Karimipour et al. 2020) focused on the two main categories of Deep Neural Network - DNN and Decision Tree - DN which came as a revolutionary influence under the deep learning model includes classifiers such as Random Forest (RM), Adaboost networks and DNN shows the available performance models based on the ten-fold cross-validation. In this work, it was suggested about a generalised ensemble deep learning-based technique for ICS-specific cyber-attack detection. A deep representation learning model is part of the suggested method, and it creates new balanced representations from the unprocessed, unbalanced dataset. In order to detect cyber-attacks, the updated representations are then applied to an ensemble deep learning system based on DNN and DT classifiers. Thereby with the help of the proposed new methodology, as an outcome of outperforming traditional classifiers shows the improved f1 score of 10% in both datasets tested and greater accuracy of 95.86% for the gas pipeline dataset and 99.67% for the secure water treatment dataset. The future work in this study is to be concentrated on improving the suggested method's accuracy and creating an extra model to distinguish between various assault types and their locations. If the proceeded majority of the system failures would be not seen and thus industrial control network would be strengthened against similar cyber-attacks.

Saghezchi et al. 2022 reported on the process of a supply chain network that interconnects man, machine and product through a revolution of industry 4.0 which is stated as the fourth step of revolution in the industry market. This input of the whole supply chain leads to challenges in the business line or the production line and even for human lives when they offer a big transformation through strategies such as the Internet of Things (IoT) and Distributed Denial-of-Service (DDoS). Thus it leads to the approach of machine learning for anomaly detection where the different kinds of attacks through DDoS could be evaluated. The trafficking of the network data could also be collected through small-scale lab test beds and information technology in which Machine Learning training

could be extracted through 45 networking bidirectional lines. The extensive simulations of supervised, non-supervised and medium-supervised eleven variants had been examined and there fourth objective result of supervised algorithms was found to be highlighted. This comes with an accurate outcome of the decision tree model limits the false positive rate to 0.001 and achieves an accuracy of 0.999.

Mousavinejad et al. 2019 elaborates the attack and detection recovery in vehicle platooning system of control where a malicious attack of cyber is highly possible as it is a platform of open and wireless communication in order to balance the sensor requirements and the data controlling command. Two kinds of attack detection had been used including a state prediction and estimated set where this algorithm found by the intersection of these two sets versus the other is the use of two recovery mechanism in which the possible cyber-attacks is examined on a time basis. The efficiency of the updated method is evaluated by the effectiveness of this application and simulation in both the detection and recovery phases. To solve the issue of detecting cyber-attacks that compromise the shared communication network and onboard sensors used in a vehicle platooning system, a novel distributed attack detection approach has been developed for future study. The simulation findings showed that the controllers in the cooperative adaptive cruise control- CCAC system can counteract the adversarial effect of an attack by accessing the safe state estimation, ensuring the string consistency of the platoon even in the presence of the assault.

Chakraborty et al. 2021 speaks about the industrial attention attack in the smart city infrastructures where the bloopers existed in this literature study leading to the multiple analysis of real-time attacks in the Internet of Things. Thus applied a machine learning approach in order to track down the Industrial attacks and it was tested recently on the (SWaT) Secure Water Treatment system. This gives the output of varying bandwidth waveforms which manipulated variable sensor readings. To accomplish this, a brand-new early detection technique was created that uses functional shape analysis (FSA) to extract features from the data and also could capture the waveform's profile. Our findings demonstrate a trade-off

between efficiency and complexity when forecasting IoT threats using functional and non-functional approaches. Numerous crucial infrastructures and systems have been compromised as a result of the rapid growth of cyber security turns into one of the major concerns facing contemporary civilization. Finally, when focusing on scenarios involving agile, real-time decision-making, the suggested functional techniques perform better than their competitors. For administrators of Industrial IoT devices and systems, these contributions collectively assist give cutting-edge threat detection capabilities and a seamless process of decision-making understanding.

Elnour et al. 2020 limelights the increasing threat of cyber activity which needs protection for the cyber security of Industrial Control Systems. Due to the present progress in the Internet of Things (IoT) and cyber activity, as well as their direct influence on several life factors, including safety, economy, and security, industrial control systems (ICS) cyber security, is becoming more and more important. To identify assaults by separating-away anomalies, the proposed cyber-attack detection system is made up of two isolation forest models that are trained independently using the normalised raw data and, respectively, a pre-processed version of the data using Principal Component Analysis (PCA). SWaT and WATI test beds had been selected as the proposed method for implementing the machine learning model approach simultaneously giving the resolution for the cyber-attack in the Industrial smart cities and factories. Feature extraction could be used to facilitate the performance of the detection approach and also to extend the vulnerable attacks to improve the capability detection of the network data in trafficking sources.

Erba et al. 2019 proposed the study on various anomalous detection using deep learning method of algorithm which would help operators to detect the challenges found in the sensor values and its cyber-physical constraints. While the research's white box attacker transforms anomalous data into regular data using an autoencoder (or a convolutional neural network), our proposed white box attacker employs an optimization strategy with a detection oracle. Our suggested methods are put into practice and assessed using two different datasets related

to the field of water distribution networks. The results demonstrate that real-time adversarial operations may drastically lower the detection algorithms' accuracy. Dealing with the BATADAL dataset, the attacker can lower the detection performance from 0.6 to 0.14. Availability attacks were given more importance compared to integrity attacks where the objective is to stimulate even detection with limited changes thereby reducing the strength of the detector. Complex sources in the future for machine learning-based attack detectors help in securing the cyber-physical systems against Evasion attacks and also assist in releasing the open source codes for dealing with the datasets available in the dataset market.

3 Methodology

3.1 Overview

In this section, the newly proposed attack detection model in ICS namely (PS-CNN) PCA Select KBest CNN has been elaborated and the overall flow is shown in Figure 4. Initially, the UNSW-NB15 dataset is loaded and the pre-processing is performed using a standard scalar process. Further to enhance the detection accuracy, the feature extraction is performed using PCA- Principal Component analysis and the feature selection process is performed using the Select KBest algorithm, selecting only the significant features. Finally, the classification of the attacks is performed using the CNN- Convolutional Neural Network classification. In the prediction phase, it detects all attack types. Finally, the proposed PS-CNN model was evaluated in terms of different performance metrics and compared with various existing models to prove its efficiency.

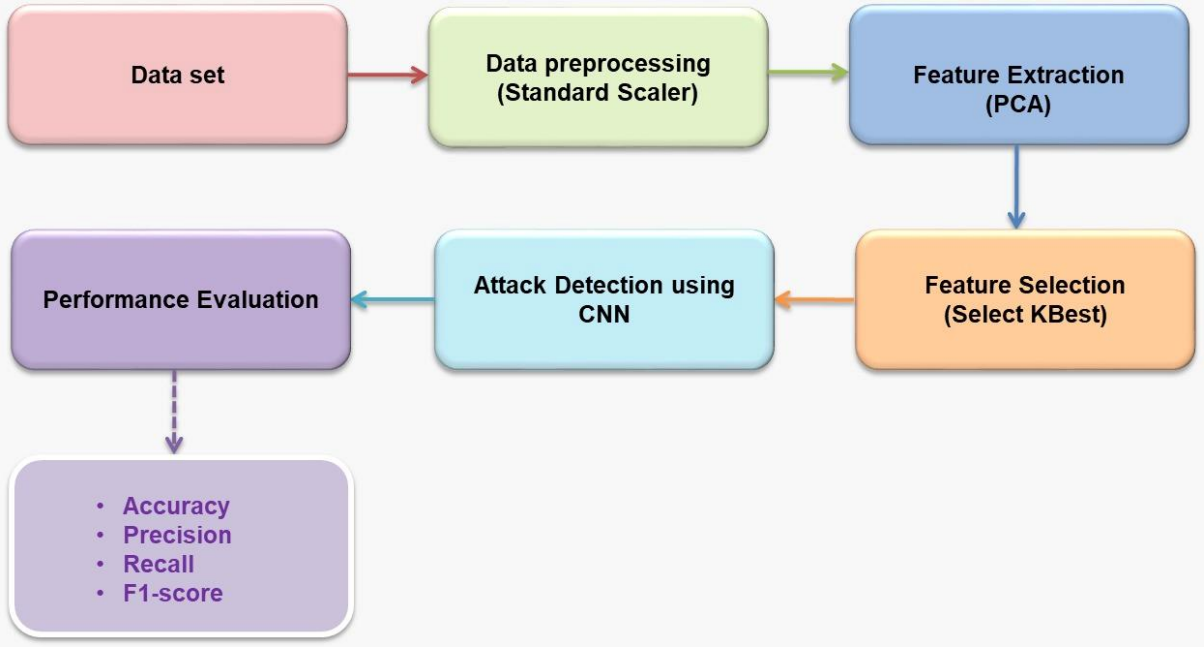

Figure 4. Proposed Glow

3.2 PCA- Principal Component Analysis for Feature Extraction

Principal Component Analysis- PCA is the well-established technique (Bro and Smilde 2014) for feature extraction used in this research. PCA is said to be the wrapper method used for feature extraction. It identifies the features around the classifier used for concentrating on the advantages of including or deleting the specific feature from the training set. PCA minimizes the number of dimensions need to classify new data and generates a set of principal components, which are termed Eigenvector pairs or Eigenvalues. It minimizes the data dimensionality by limiting the attention to feature space directions in which the greatest variance exists. The ratio of total variance concerning features is proportional to their Eigenvalue. Let's as me $(\mathrm{x_t})$, $\mathrm{t} = 1,2, \ldots, \mathrm{N}$ are stochastic with n-dimensional input data records with (μ) mean value. $\mu = \frac{1}{N}\sum_{t=1}^{N} X^t$

Equation 1. (μ) mean value

$C\ is$ covariance matrix is determined as,

$$C = \frac{1}{N}\sum_{t=1}^{N} (X_t - \mu)\,.\,(X_t - \mu)$$

Equation 2. Covariance matrix

PCA can solve the covariance matrix complexity of the Eigenvalue,

$$Cv_i = \lambda_i v_i$$

Equation 3. Eigenvalue

Here, Eigen values are $\lambda_i\ (i = 1, 2, \ldots, n)$ and corresponding Eigenvectors are $v_i\ (i = 1, 2, \ldots, n)$.

3.3 Select KBest for Feature Selection

The feature selection process is used for selecting the significant features from a dataset which contributed to the target variable. For the target variable, the best predictors are selected. Select KBest module classes are used for dimensionality reduction and feature selection for improving the accuracy scores or improvising the performance on high-dimensional datasets. The proposed Select KBest method can minimize the overfitting through lesser redundant data termed as decision-making in the least possibility concerning redundant noise or data. It further enhances accuracy, through lesser misleading information. Training time is lesser for the Select KBest approach and thus algorithm can train faster.

3.4 Classification using CNN

Convolutional Neural Network- CNN-based deep learning approach comprised of convolutional layers. The feature maps are extracted through convolutional layers from the input image using various numbers of kernels. Followed by pooling layers which reduce these dimensions. Different pooling layers are also presented such as average pooling and max pooling. Between the convolutional and fully connected layers, subsampling layers are presented. CNN can be applied automatically in datasets where relatively greater numbers of parameters and nodes are trained.

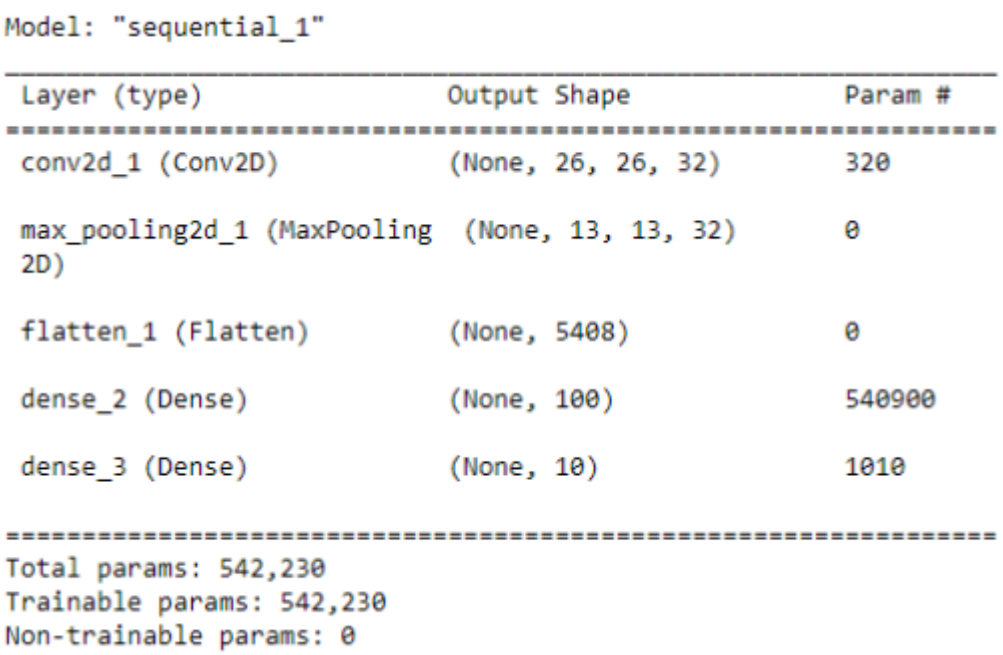

Figure 5. CNN Parameters

3.4.1 Input layer

From the pre-processed data, the selected features are determined as convolutional neural networks input vectors, which are termed as,

$$X = [x_1, x_2, x_3 \dots x_k]$$

Equation 4. CNN input vector

From Eq.5, no. of features per window is k after measurement. To speed up the model convergence, the Min-Max Normalization formula is used in which the values in every dimension of data are transformed linearly and normalized to [0,1] range,

$$x = \frac{x - \min}{\max - \min}$$

Equation 5. Min-Max Normalization

From Eq.(5), every column's minimum is $\min$, whereas the maximum value is $\max$.

3.4.2 Convolutional layer

The j th feature map output on i th unit of convolutional layer l is determined as,

$$x_i^{l,j} = \sigma\left(b_j + \sum_{a=1}^{m} w_a^j \; x_{i+a-1}^{l-1,j}\right)$$

Equation 6. Feature map output

From Eq. (6), for j the feature map, bias term is b_j, kernel size is m, j th feature map weight is w_a^j , σ is the activation function and filter index is a th. Here ReLu activation function has used, with faster training speed.

3.4.3 Max-pooling layer

For closer output of $x_i^{l,j}$, the pooling layer expressed the aggregation statistics, can minimize the sensitivity and dimension of output and attains the feature preservation based scale-invariant. Max pooling operation is used in this research. The max pooling layer divides the convolution layer output features into multiple partitions and identifies the maximum in every partition. The max pooling layer output is,

$$x_i^{l,j} = \max_{n=1}^{r}\left(x_{(i-1)*T_n}^{l-1,j}\right)$$

Equation 7. Max pooling layer output

From Eq.(7), the pooling size is n and pooling stride is T

3.4.4 Model Training

CNN has comprised of several convolutional layers, fully connected layers and pooling layers. The input data has mapped into the feature space of the hidden layer after the pooling, convolutional and activation function's operation. For final classification performance, the fully connected layer is combined finally with various local structural features that are learned from the below layer. Here in this proposed research, only one pair of max-pooling and convolutional layers is considered, further 2D data is flattened to 1D data and with a fully connected layer the neural network is completed. The soft-max classifier is used as the model's top layer for recognising the transportation modes,

$$f(x) = \mathrm{argmax}_c\left(\frac{e^{x^{l-1} w_j}}{\sum_{n=1}^{N} e^{x^{l-1} w_n}}\right)$$

Equation 8. Soft-max classifier

From Eq. (8), the class label is c, the sample feature is x, the index layer is l and the number of the class is N. concerning Eq. (3.6) to (3.8), forward propagation has been processed. From the input layer, the information is forward propagated, to the output layer through the hidden layer and attains the whole network output. The iteration of forward propagation attains the network error value. The cost function of cross-entropy has approved to measure the error value is,

$$L(y) = -\frac{1}{n}\sum_{x}[y \ln a + (1-y)\ln(1-a)]$$

Equation 9. Cost function of cross-entropy

From Eq. (9), the sample is x, the number of training samples is n, the network predicted value is a and the actual value is y. A gradient descent algorithm has been adopted during the training process for parameter adjustment (bias b and weight w) layer by layer concerning error backward propagations.

4 Results and Discussion

4.1 Dataset Description

The proposed feature selection-based RF classifier method is used to detect the intrusion on the UNSW-NB15 dataset. The UNSW-NB15 dataset raw network packets have been generated by UNSW Canberra cyber range lab for synthetic contemporary behaviour attacks and generating activities of normal and suspicious. The UNSW-NB15 dataset(Moustafa and Slay 2014) was generated using the IXIA PerfectStorm tool in the cyber range lab of ACCS- Australian Centre for cyber security for creating realistic modern normal activities hybridization and from network traffic, synthetic contemporary attack behaviours focused. A tcpdump tool is utilized to record network traffic of 100GB. Bro-IDS and Argus's tools were utilized and for extracting features, 12 models have been developed using Table 1-5(Moustafa and Slay 2016).

Table 1: Flow features

No.	Name	Description
1	*Srcip*	Source IP address
2	*Sport*	Source port number
3	*Dstip*	Destination IP address
4	*Dsport*	Destination port number
5	*Proto*	Protocol type (such as TC, UDP)

Table 2: Standard features

6	*state*	Indicates to the state and its dependent protocol (such as ACC, CLO and CON).
7	*dur*	Record total duration
8	*sbytes*	Source to destination bytes
9	*dbytes*	Destination to source bytes
10	*sttl*	Source to destination time to live
11	*dttl*	Destination to source time to live
12	*sloss*	Source packets retransmitted or dropped
13	*dloss*	Destination packets retransmitted or dropped
14	*service*	Such as http, ftp, smtp, ssh, dns and ftp-data.
15	*sload*	Source bits per second
16	*dload*	Destination bits per second
17	*spkts*	Source to destination packet count
18	*dpkts*	Destination to source packet count

Table 3: Content features

19	*swin*	Source TCP window advertisement value
20	*dwin*	Destination TCP window advertisement value
21	*stcpb*	Source TCP base sequence number
22	*dtcpb*	Destination TCP base sequence number
23	*smeansz*	Mean of the flow packet size transmitted by the src
24	*dmeansz*	Mean of the flow packet size transmitted by the dst
25	*trans_depth*	Represents the pipelined depth into the connection of http request/response transaction
26	*res_bdy_len*	Actual uncompressed content size of the data transferred from the server's http service

Table 4: Time features

27	*sjit*	Source jitter (mSec)
28	*djit*	Destination jitter (mSec)
29	*stime*	record start time
30	*ltime*	record last time
31	*sintpkt*	Source interpacket arrival time (mSec)
32	*dintpkt*	Destination interpacket arrival time (mSec)
33	*tcprtt*	TCP connection setup round-trip time, the sum of 'svnack' and 'ackdat'

Table 5: Additional features

36	*is_sm_ips_ports*	If *srcip* (1) equals to *dstip* (3) and *sport* (2) equals to *dsport* (4), this variable assigns to 1 otherwise 0
37	*ct_state_ttl*	No. for each *state* (6) according to specific range of values of *sttl* (10) and *dttl* (11)
38	*ct_flw_http_mthd*	No. of flows that has methods such as Get and Post in http service
39	*is_ftp_login*	If the ftp session is accessed by user and password then 1 else 0
40	*ct_ftp_cmd*	No of flows that has a command in ftp session
41	*ct_srv_src*	No. of records that contain the same *service* (14) and *srcip* (1) in 100 records according to the *ltime* (26)
42	*ct_srv_dst*	No. of records that contain the same *service* (14) and dstip (3) in 100 records according to the *ltime* (26)
43	*ct_dst_ltm*	No. of records of the same *dstip* (3) in 100 records according to the *ltime* (26)
44	*ct_src_ ltm*	No. of records of the *srcip* (1) in 100 records according to the *ltime* (26)
45	*ct_src_dport_ltm*	No of records of the same *srcip* (1) and the *dsport* (4) in 100 records according to the *ltime* (26)
46	*ct_dst_sport_ltm*	No of records of the same *dstip* (3) and the *sport* (2) in 100 records according to the *ltime* (26)
47	*ct_dst_src_ltm*	No of records of the same *srcip* (1) and the *dstip* (3) in in 100 records according to the *ltime* (26)

Table 6: Dataset Distribution

Category	Training set	Testing set
Normal	56,000	37,000
Analysis	2,000	677
Backdoor	1,746	583
DoS	12,264	4089
Exploits	33,393	11,132
Fuzzers	18,184	6,062
Generic	40,000	18,871
Reconnaissance	10,491	3,496
Shellcode	1,133	378
Worms	130	44
Total Records	175,341	82,332

Table 6, explains the generation of training and testing sets from the selected dataset UNSW-NB15 dataset. A part of the dataset observed and records have been divided with a 60:40 per cent ratio approximately as training and testing set. In attaining the IDS evaluation authenticity, no redundant records between the training and testing sets are recorded.

4.2 Performance Analysis

The following are the results of the proposed PS-CNN used for detecting attacks in ICS.

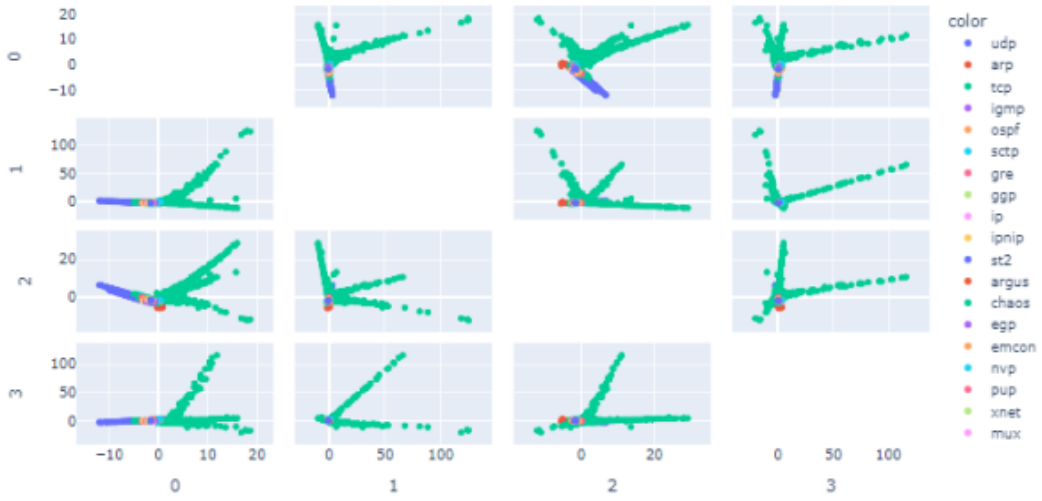

Figure 6. Feature extraction results

The feature extraction process for the proposed system is shown in Figure 6. The colour variation for the categorized features enables a better understanding of the feature extraction step. In general, the feature extraction process aims to decrease the number of features in the provided dataset by generating new features from the prevailing features. Such new reduced features must be able to summarize the data comprised in the raw dataset. With the use of the employed regularisation phenomenon by the proposed method, the issue associated with overfitting can be resolved. Further, the proposed method could be able to adapt accuracy improvements, speed, and improved visualisation of the data.

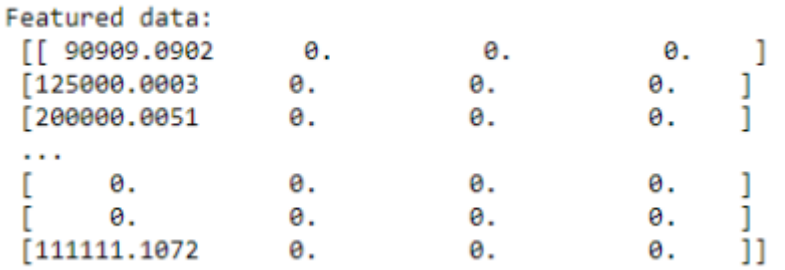

Figure 7. FEATURE SELECTION – Univariate Selection (SelectKBest)

Figure 7 denotes the univariate selection values of the K Best feature selection model. The univariate selection of the proposed system works effectively by choosing the best features in accordance with the univariate statistical tests. It could be observed as the pre-processing step for the estimator. The select K Best eliminates all the irrelevant features.

Table 7: Performance analysis

Performance metrics	Proposed
Accuracy	0.99971
F1-score	0.99716
Precision	0.99716
Recall	0.99716
RMSE	1.764

The above table denotes the values of the performance metrics such as accuracy, F1 score, precision, recall and RMSE values of the proposed system. The accuracy is the ratio between the number of correct predictions to the total number of predictions. F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. Precision is the ratio of correctly predicted positive examples divided by the total number of positive examples that were predicted. The recall is a metric that quantifies the number of correct positive predictions made out of all positive predictions that could have been made. RMSE explains how accurate our predictions are and, what is the amount of deviation from the actual values. From the table, it is observed that the accuracy is 99.97%, the F score was 99.71%, the precision is found to be 99.71%, the recall is found to be 99.71% and RMSE is 1.764. These values prove the efficiency of the proposed system.v

Table 8: Comparative analysis

Methods	**Accuracy**
NO	95.98
SVM	95.51
D-CNN	96.50
RNN	95.83
TABOR	95.79
DIF	98.03
ID CNN-IF	92.01
Proposed (PS-CNN)	99.71

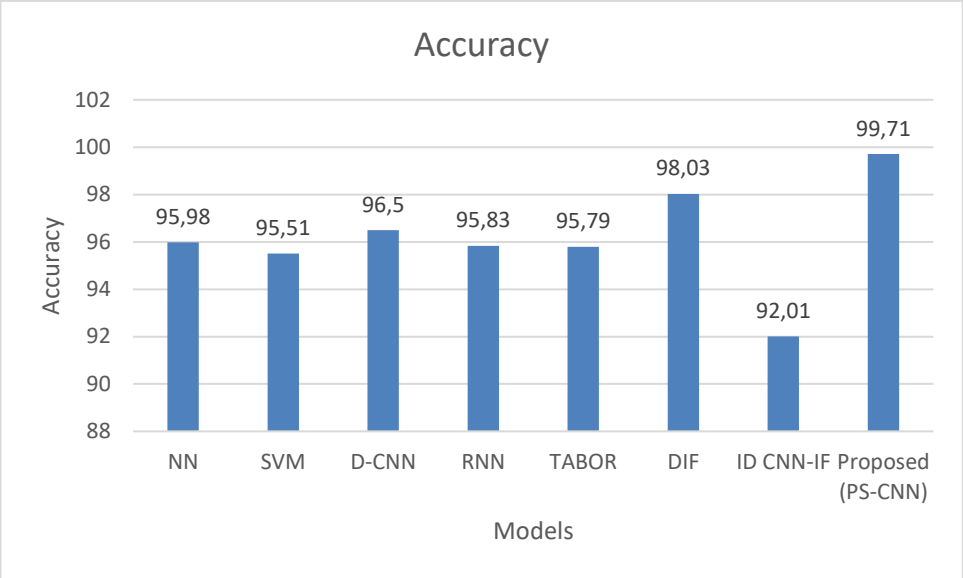

Figure 8. Accuracy comparison

The accuracy of the proposed system in comparison with the other existing models like NN, D CNN, RNN, SVM, TABOR, DIF, and ID CNN IF are listed in the above table. Here it is observed that the proposed system outperforms the other models with an accuracy value of 99.71. This high performance is achieved predominantly through the employment of the prominent function of the data pre-processing, feature extraction, feature selection and attack detection of the proposed PS-CNN- PCA Select KBest CNN method system.

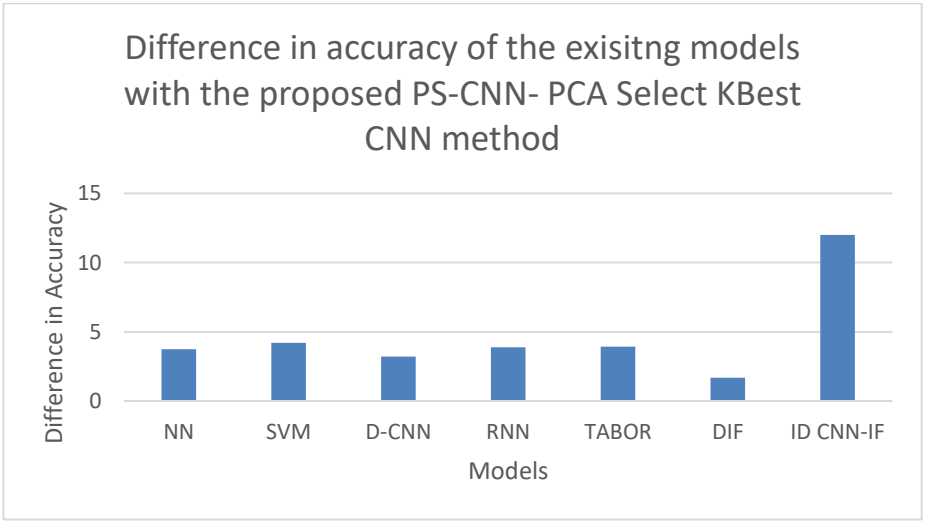

Figure 9. Comparative Analysis

It is also viewed in the table that NN attains 95.98 accuracy which is 3.73% lower than the accuracy of the proposed system. Accordingly, the accuracy of the Support Vector machine has been observed to be 95.51 which is 4.2 % lower than our proposed system. Meanwhile, the accuracy of the D-CNN has been observed to be 96.51 which is 3.21 % lower than our proposed system. Likewise, the accuracy of the RNN has been observed to be 95.83 which is 3.88 % lower than our proposed system. Similarly, the accuracy of the TABOR has been observed to be 95.79 which is 3.92 % lower than our proposed system. The accuracy of the DIF has been observed to be 98.3 which is 1.68% lower than our proposed system. Finally, it is also observed that the accuracy of the ID-CNN-IF has been observed to be 92.1 which is 12 % lower than our proposed system.

Despite the approaches suggested by previous works also having good accuracy values and low false alarm rates when compared with the CNN IF model, it is observed that the latter model can be able to identify a wide range of attack scenarios. Our proposed system could also be able to detect an extensive range of attacks scenario when compared with all other models. Additionally, the computational complexity of the existing approaches is very low when compared with the proposed model and hence the detection process is also high with the proposed method. The dynamic behaviour of the proposed method or extremely high in lowering the computational complexity and hence reliable detection of the attacks can be possible with the proposed scenario in normal operation and complex operations. Future work is associated with further enhancement of the performance of the proposed method especially in terms of false alarm rate. Accordingly, the proposed approach can be also tested with other data searches in future.

5 Conclusion and Future Work

This research has focused on designing an efficient attack detection model for cyber-attack detection in ICS. Three objectives have been framed to achieve this target. i.e., an effective Attack detection model namely (PS-CNN) PCA Select KBest CNN designed for detecting attacks in Industrial Control System with better accuracy in detection and lesser false positives. Further the informative features identified from the selected UNSW-NB15 dataset using the Principle component Analysis for feature extraction and Select KBest approach for the feature selection process. Finally, malicious attacks in ICS with higher accuracy in attack detection are classified using Convolutional Neural Network- CNN. These three objectives have been successfully designed with a robust attack detection model.

High-dimensional data significantly impact identifying the attack vectors with better accuracy. Hence, a feature selection technique based on Select KBest was designed to enhance the performance of the attack detection approaches with reduced time complexity and better accuracy. The detection capability of the classification algorithm was improved via the CNN algorithm. The results obtained confirmed that the developed system can be applied to any ICS infrastructure as the false positives were less for two different ICS environments. Anomalies were also identified in a lesser time. From the extensive experiments and results obtained, it is worth noting that the proposed attack detection model is not limited to a single dataset. Our models were validated for different ICS infrastructures. However, this thesis work is limited to detecting cyber-attacks in ICS operations. Identification of faulty components was not in the scope of this thesis work which is considered to be a setback. To ensure the continuous operations of the system, it is necessary to differentiate between attacks and fault error rates. Beyond addressing these shortcomings, this research can be extended in many possible ways such as the Identification of anomalies for the faulty system, and the Integration of machine learning and deep learning-based attack detection module with an open-source IDS.

References

Abana, M. A., et al. (2016). "Coverage and rate analysis in heterogeneous cloud radio access networks with device-to-device communication." IEEE Access **4**: 2357-2370.

Al-Abassi, A., et al. (2020). "An ensemble deep learning-based cyber-attack detection in the industrial control system." IEEE Access **8**: 83965-83973.

Alladi, T., et al. (2020). "Industrial control systems: Cyberattack trends and countermeasures." Computer Communications **155**: 1-8.

Bro, R. and A. K. Smilde (2014). "Principal component analysis." Analytical methods **6**(9): 2812-2831.

Chakraborty, S., et al. (2021). "Machine learning for automated industrial IoT attack detection: an efficiency-complexity trade-off." ACM Transactions on Management Information System (TMIS) **12**(4): 1-28.

Elnour, M., et al. (2020). "A dual-isolation-forests-based attack detection framework for industrial control systems." IEEE Access **8**: 36639-36651.

Erba, A., et al. (2019). "Real-time evasion attacks with physical constraints on deep learning-based anomaly detectors in industrial control systems." arXiv preprint arXiv:1907.07487.

Kravchik, M., et al. (2021). Poisoning attacks on cyber attack detectors for industrial control systems. Proceedings of the 36th Annual ACM Symposium on Applied Computing.

Kravchik, M. and A. Shabtai (2021). "Efficient cyber attack detection in industrial control systems using lightweight neural networks and pca." IEEE Transactions on Dependable and Secure Computing.

Lai, Y., et al. (2019). "Industrial anomaly detection and attack classification method based on convolutional neural network." Security and Communication Networks **2019**.

Li, B., et al. (2020). "DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems." IEEE Transactions on Industrial Informatics **17**(8): 5615-5624.

Lu, K.-D., et al. (2021). "Evolutionary deep belief network for cyber-attack detection in industrial automation and control system." IEEE Transactions on Industrial Informatics **17**(11): 7618-7627.

Mousavinejad, E., et al. (2019). "Distributed cyber attacks detection and recovery mechanism for vehicle platooning." IEEE Transactions on Intelligent Transportation Systems **21**(9): 3821-3834.

Moustafa, N. and J. Slay (2014). Unsw-nb15 dataset for network intrusion detection systems.

Moustafa, N. and J. Slay (2016). "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set." Information Security Journal: A Global Perspective **25**(1-3): 18-31.

Noorizadeh, M., et al. (2021). "A cyber-security methodology for a cyber-physical industrial control system testbed." IEEE Access **9**: 16239-16253.

Saghezchi, F. B., et al. (2022). "Machine learning for DDoS attack detection in industry 4.0 CPPSs." Electronics **11**(4): 602.

Zhang, D., et al. (2021). "A survey on attack detection, estimation and control of industrial cyber–physical systems." ISA transactions **116**: 1-16.

Zyad, E., et al. (2019). Improve R2L attack detection using trimmed PCA. 2019 International Conference on Advanced Communication Technologies and Networking (CommNet), IEEE.