The ISC Int'l Journal of Information Security
ISeCure

November 2023, Volume 15, Number 3 (pp. 43–50)
http://www.isecure-journal.org

Presented at the ISCISC'2023 in Tehran, Iran.

A Semi-Supervised IDS for Cyber-Physical Systems Using a Deep Learning Approach **

Amirhosein Salehi [1,*], Siavash Ahmadi [2], and Mohammad Reza Aref [1]

[1] *Information Systems and Security Lab, Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran*
[2] *Electronics Research Institute, Sharif University of Technology, Tehran, Iran*

ARTICLE INFO.

Keywords:
Autoencoder, Cyber-Attack, Industrial Control Systems, Intrusion Detection System, Deep Learning

Type:
Research Article

doi:
10.22042/isecure.2023.181544

ABSTRACT

Industrial control systems are widely used in industrial sectors and critical infrastructures to monitor and control industrial processes. Recently, the security of industrial control systems has attracted a lot of attention, because these systems are now increasingly interacting with the Internet. Classic systems are suffering from many security problems and with the expansion of Internet connectivity, they are now exposed to new types of threats and cyber-attacks. Addressing this, intrusion detection technology is one of the most important security solutions that is used in industrial control systems to identify potential attacks and malicious activities. In this paper, we propose Stacked Autoencoder-Deep Neural Network (SAE-DNN), as a semi-supervised Intrusion Detection System (IDS) with appropriate performance and applicability on a wide range of Cyber-Physical Systems (CPSs). The proposed approach comprises a stacked autoencoder, a deep learning-based feature extractor, helping us with a low dimension and low noise representation of data. In addition, our system includes a deep neural network (DNN)-based classifier, which is used to detect anomalies with a high detection rate and low false positive rate in a real-time process. The SAE-DNN's performance is evaluated on the WADI dataset, which is a real testbed for a water distribution system. The results indicate the superior performance of our approach over existing supervised and unsupervised methods while using a few percentages of labeled data.

1 Introduction

In today's world, the control of physical systems is based on methods where all processes can be connected and controlled by communication links. These systems, which result from the integration and coordination of cyber and physical components are called Cyber-Physical Systems.

* Corresponding author.
**The ISCISC'2023 program committee effort is highly acknowledged for reviewing this paper.
Email addresses: ah.salehi@ee.sharif.edu, s.ahmadi@sharif.edu, aref@sharif.edu
ISSN: 2008-2045

CPSs are found in critical infrastructures such as water distribution, energy, and transportation and consist of a physical process controlled by an Industrial Control System (ICS). In a CPS, a set of sensors measure process variables, such as temperature, flow rate, level, etc., from the physical process and send these values to the controllers through communication channels. Based on these values, the controller makes decisions and initiates actions in the physical process through the actuators [1]. These controllers, sensors, and actuators make a large amount of data that can be used for continuously monitoring the system processes and detecting anomalies.

CPSs have many open security vulnerabilities due to a lack of encryption and authentication in their communication protocols. Therefore, these systems are susceptible to many types of attacks and intrusions, such as hijacking the communication links, spoofing network messages, modification of sensor data, packet sniffing for using in subsequent attacks or changing some control commands to break the normal state of the system. In addition, with the integration of OT and IT systems, which connects the CPSs to the Internet, these vulnerabilities may lead to many new other attack types that will cause harmful damage to the whole system. With the advent of these threats, the existence of such systems that can continuously monitor all processes and communication links is crucial.

However, existing solutions have failed to provide plausible security for OT systems. The design of IT-based IDSs Makes them incapable of monitoring the measurement level activities of physical processes, so they cannot be Exploited to detect multi-stage attacks. Furthermore, the architecture of existing solutions is based on a limited set of instructions (rule-based systems) given to the system to detect and block abnormal traffic. This limitation renders these solutions vulnerable to unknown threats like zero-day attacks. As a result, any subversive behavior that does not contradict their instructions will be hidden from their view. As a solution, industrial IDS with capabilities such as continuous analysis and monitoring of the industry's internal network (OT), learning from previous and existing data, as well as proper performance against attacks of foreign origin, can provide high-level security for ICSs.

Regarding the detection methods, IDSs can be categorized into four baseline approaches as follows:

1. Statistical-based IDSs: these systems process events or network traffic by statistical algorithms (such as parametric and non-parametric methods, time series analysis, Markov chains, etc.) to check whether a piece of data matches a given statistical model or not, which confirms the presence of intrusion. When an event occurs during anomaly detection, it is assigned an anomaly score, which is trained by comparing the current and the previous statistical profile. The anomaly score shows the degree of irregularity of the specific event, and if the anomaly score is higher than a certain threshold, the system generates an alarm.

2. Rule-based IDSs: in these systems, the traffic packets are screened by a decision engine based on a database of intrusion detection rules it is fed. These packets are labeled as anomalies and discarded if an error is detected by any of the rules. These systems have a high detection accuracy; but very low efficiency in detecting unknown and novel attacks.

3. Signature-based IDSs: a signature is a pattern or string associated with a known attack or threat. A signature-based intrusion detection system is a process of comparing patterns with recorded events to identify possible intrusions. This system is also known as knowledge-based detection or misuse-based detection due to the use of knowledge accumulated by specific attacks and system vulnerabilities. These systems are also vulnerable to unknown and new attacks since they are built on a limited set of attack signatures.

4. Deep learning-based IDSs: After observing the vulnerability of traditional intrusion detection systems like misuse-based detection methods, researchers turned to the design of new systems based on machine learning algorithms. In these systems, attack detection is performed based on the output of a classifier trained by a dataset, including normal and abnormal network traffic samples. To aim this, DNNs can profit the scientists in this field, as they obviate the need for feature selection and provide a high classification accuracy. One of the advantages of deep learning-based approaches compared to the classical machine earning-based systems is higher detection accuracy while having a lower rate of false positive alerts, which has attracted researchers to design new systems based on this approach. In the last five years, they have designed different systems based on these algorithms. In general, the design phases of a learning-based system are divided into three categories: 1) data collection, 2) selecting and extracting features, and 3) a decision-making engine.

1.1 Our Contributions

In this paper, we propose SAE-DNN, a noise-resistant semi-supervised method for cyber-physical systems that uses a stacked autoencoder and a deep neural network classifier to detect malicious data. Due to the lack of sufficient labeled data in practical cases,

a semi-supervised approach would be helpful that mainly uses unlabeled data while using a low percentage of labeled data for the training phase. Since the WADI dataset is highly dimensional and has 126 features and many of these features are useless, the proposed approach should have a dimension reduction and feature extraction phase before any classifier or detector. Hence, our proposed method comprises a stacked autoencoder that has both of these properties. In addition, the stacked autoencoder acts as a noise reducer phase that is so helpful due to the noise of measurements when monitoring the system and collecting data. To classify the data as normal or malicious, we propose a four layers-deep neural network that detects the abnormal samples with a high detection rate while having a low false positive rate. Also, to address the imbalanced dataset issue, we use SMOTE technique which helps us with a significantly improved detection rate. Briefly, our main contributions are:

- Proposing a semi-supervised approach with good performance and applicable to a wide range of CPSs.
- Using deep auto-encoder as a feature extractor that also helps us with dimension and noise reductions.
- Using a four-layer (two hidden layers and two layers of input and output) deep neural network to classify the data into normal and malicious.
- Fine-tuning the network using a few percent of labeled data, which significantly improves the detection rate of our model.
- Using SMOTE oversampling technique to address the imbalanced dataset issue.

1.2 Paper Organization

The remainder of this paper is organized as follows. In Section 2, we explain some related works. Section 3 introduces the preliminaries on which our proposed approach is based. Section 4 explains our proposed architecture by describing the stacked autoencoder architecture; and deep neural network. In Section 5, we present and discuss the achieved results. Finally, Section 6 concludes the paper and proposes possible future extensions to this work.

2 Related Work

An anomaly detection method based on machine learning has been widely studied and achieved good results in recent years. Due to the inherent lack of labeled anomaly data for training supervised algorithms, anomaly detection methods are mostly based on unsupervised methods. In this section, we briefly describe the approaches that used the WADI dataset to train and test their algorithms. These approaches can be generally divided into two categories: 1) GAN-based and 2) basic classic methods. Dan Li *et al.* [2] proposed an unsupervised multivariate anomaly detection method based on Generative Adversarial Networks (GANs), using the Long-Short-Term-Memory Recurrent Neural Networks (LSTM-RNN) as the base model (namely, the generator and discriminator) in the GAN framework to capture the temporal correlation of time series distributions.

Araujo-Filho *et al.* [3] proposed FID-GAN, a novel fog-based unsupervised IDS for CPSs using GANs. To achieve higher detection rates, the proposed architecture computes a reconstruction loss based on the reconstruction of data samples mapped to the latent space. They also addressed the problem of latency-constrained applications by training an encoder that accelerates the reconstruction loss computation. ZHANG *et al.* [4] proposed Transferred Generating Adversarial Network-Intrusion Detection System (TGAN-IDS) based on dual generative adversarial networks. A Deep Convolutional Generative Adversarial Network (DCGAN) was adopted to train a generator and was transferred to the generator of TGAN. A pre-training model named PreD was built based on Convolutional Neural Network (CNN) to do binary classification and was transferred to the discriminator of TGAN. They also introduced a reconstruction loss function into the target function of TGAN to suppress the deterioration of normal sample detection ability during adversarial training of TGAN.

Kayan *et al.* [5] Proposed AnoML, which is an end-to-end data science pipeline that allows the integration of multiple wireless communication protocols, anomaly detection algorithms, and deployment to the edge, fog, and cloud platforms with minimal user interaction. They also evaluated the pipeline with two anomaly detection datasets while comparing the efficiency of several machine learning algorithms within different nodes. ELNOUR *et al.* [6] Proposed a novel semi-supervised Dual Isolation Forests-based (DIF) attack detection system that has been developed using the normal process operation data only and is composed of two isolation forest models that are trained independently using the normalized raw data and a pre-processed version of the data using Principal Component Analysis (PCA), respectively, to detect attacks by separating-away anomalies.

Alsaedi *et al.* [7] proposed a framework, named Unsupervised Misbehavior Detection (USMD), comprising a deep neural network that learns about a system's expected behavior from data-driven representations. USMD can identify in real-time the at-

tacks on CPSs by using the long-short-term memory and attention method for multi-sensor data. Shahid *et al.* [8] compared two types of IDSs, namely design-based and data-based approaches, and then applied some machine learning-based system-modeling techniques to the dataset. Design-based approaches require domain expertise and are not scalable. On the other hand, data-based approaches suffer from the lack of real-world datasets available for specific critical physical processes. They also proposed an operational invariants-based attack detection technique using the system design parameters. Kabir *et al.* [9] applied and compared several supervised machine learning methods, such as k-nearest neighbors (kNN), Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) using WADI test dataset.

3 Preliminaries

3.1 Autoencoder

An autoencoder is an unsupervised artificial neural network, as the output is the reconstruction of the input itself [7]. In particular, autoencoders are trained to learn a mapping function from the input to itself, defined as

$$\hat{X} \cong \phi(\varphi(X)) \tag{1}$$

where X is the input data, ϕ is an encoder that maps the input X into some latent variables Z, and is a decoder that maps Z back into the input space as $\hat{X}$. The training objective is to train φ and ϕ to minimize the reconstruction error, which is the difference between the original input X and the reconstructed output $\hat{X}$. Thus, an autoencoder can be seen as a solution to the following optimization problems:

$$\min \|X - \phi(\varphi(X))\|^2 \tag{2}$$

where $\|.\|^2$ denotes the l_2-norm. Figure 1 illustrates an example of an autoencoder.

3.2 Deep Neural Network

In classical machine learning, the important features of input are manually selected, and the system automatically learns to map the feature space into the output variable. There are multiple levels of features in deep learning. These features are automatically discovered and composed together at various levels to produce outputs. Each level represents abstract features discovered from the features presented in the previous level [10].

3.3 Evaluation Metrics

There are some metrics to evaluate the performance of IDSs, which we briefly explain as follows:

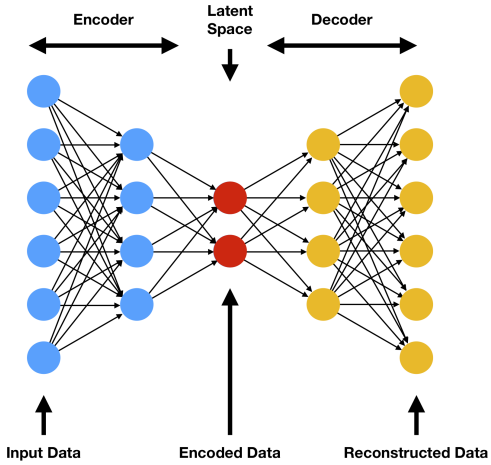

Figure 1. An example of an autoencoder [12]

$$Precision = \frac{TP}{(TP + FP)} \times 100$$

$$Recall = \frac{TP}{(TP + FN)} \times 100$$

$$F1score = 2 \times \frac{Precision \times Recall}{(Precision + Recall)} \times 100$$

$$FPR = \frac{FP}{(TN + FP)} \times 100$$

$$Accuracy = \frac{TP + FN}{(TP + FP + TN + FN)} \times 100$$

where TP and FP represent the number of positive samples that are correctly classified and misclassified, respectively. TN and FN also represent the number of negative samples that are correctly classified and misclassified, respectively. We evaluate these parameters as they are the core elements that determine the efficiency of an anomaly detection algorithm. Accuracy determines the overall correct prediction ratio. Precision demonstrates how many of the predicted anomalies are really anomalies. Recall demonstrates how many of anomalies are detected. F1˙Score is an evaluation metric that considers class distribution.

4 The Proposed Approach

In this section, we explain our proposed approach. The system architecture of our method comprises a stacked autoencoder as a feature extractor and a DNN classifier to detect anomalies. The stacked autoencoder takes normal samples as input and trains itself to represent a low-dimension and low-noise representation of input data. Then a total model is created by cascading the weights of encoders weights of trained stacked autoencoder to a four-layer deep learning classifier.

Afterward, the labeled data, a few percent of the whole testing dataset and including normal and under-attack samples, will be given into the total model to

accomplish the training process. Meanwhile, the first part of the model (weights of trained stacked autoencoder) is set to untrainable. After the second training phase is completed, we set the first part of the model to trainable mode, the second part (classifier) to untrainable mode, and then fine-tune the first part of our model with labeled data, which means retraining the first part with a low learning rate. Here we explain the architecture of our stacked autoencoder, the DNN classifier, and the total model that is cascaded from these two parts.

4.1 Stacked Autoencoder

The stacked autoencoder we created consists of two separate autoencoders. This model works as follows:

- The first autoencoder is trained using input data (normal samples).
- Encoded data from the first autoencoder will be given as input to the second autoencoder for training the model.
- After the training is completed, the encoder layers of the first and second autoencoder are cascaded and create the first part of the total model (feature extractor).

The first autoencoder consists of input and output layers of 127 neurons (126 features) and an encoded layer of 64 neurons. The second autoencoder consists of input and output layers of 64 neurons and an encoded layer of 32 neurons. Figure 2 shows the architecture of the proposed feature extractor. We used adam optimizer and set the learning rate to 1e-3.

4.2 DNN Classifier

We used a four-layer classifier (two hidden layers) which consists of an input layer of 32 neurons, an output layer with one neuron (binary classifier), and two hidden layers with 32 and 16 features, respectively. We also used sigmoid as an activation function. Figure 3 illustrates the architecture of our proposed DNN classifier. As shown in Figure 3, the classifier consists of a four-layer input of 32 neurons, two hidden layers of 32 and 16 layers respectively, and a binary output layer.

4.3 Fine-Tuning

In the first step, we trained our stacked autoencoder using unlabeled normal data. After that, we made the total model and trained the DNN classifier using labeled data. After completion of these two training phases, we designed a fine-tuning step, which is retraining the feature extractor leveraging labeled data used to train the classifier. The learning rate of the fine-tuning step should be lower than the learning rate of previous steps so as not to disturb the convergence of the feature extractor. We set the learning rate to 1e-4 and set the number of epochs to 60. All experiments were conducted on an Intel(R) Core(TM) i5-6200U CPU @ 2.30GHz with 8GB of RAM.

5 Results and Discussion

In this section, we evaluate our model and compare it with unsupervised and supervised algorithms, which used the WADI dataset to implement their models. First, we describe the dataset used to train and test our model.

5.1 Datasets

We used WaterDisribution (WADI) dataset to train and evaluate our proposed model. Water distribution networks are often geographically spread and require automatic control to operate. Automation makes the water distribution network vulnerable to cyber-physical attacks [11]. The WADI dataset represents an extension of a Secure Water Treatment (SWAT) system by considering a complete and realistic water treatment, storage, and distribution network. It contains 1,209,610 data patterns with 126 features by collecting 126 sensors and actuators data from a water distribution testbed that had non-stop run for 16 days while being attacked during the last two days [5]. The first 14 days of the process represent the normal behavior of the system and the data collected during this period is used to generate the training dataset. The last 2 days of the process are under attack and the collected data is labeled and used to generate the test dataset. We divide the test dataset into two separate datasets. Then we use one of them for classifying and fine-tuning the system and the other one for testing the total model. The goal of an attacker is to manipulate the normal operations of the plant by changing the reading of values of sensors or actuators. It is assumed that the attacker has remote access to the SCADA system of WADI and has general knowledge about how the system works. Various experiments have been conducted on the WADI system to investigate cyber-attacks and respective system responses. In total, 15 attacks have been inserted into WADI [2].

5.2 Evaluation

To train our models, we have used the normal samples and 1%, 2%, 3%, and 10% of the labeled dataset as labeled samples, respectively. Table 1 illustrates the comparison of the results of our proposed approach with some of the best previous results of unsupervised and supervised algorithms. As the table shows, for example using only 1% of labeled data, the results

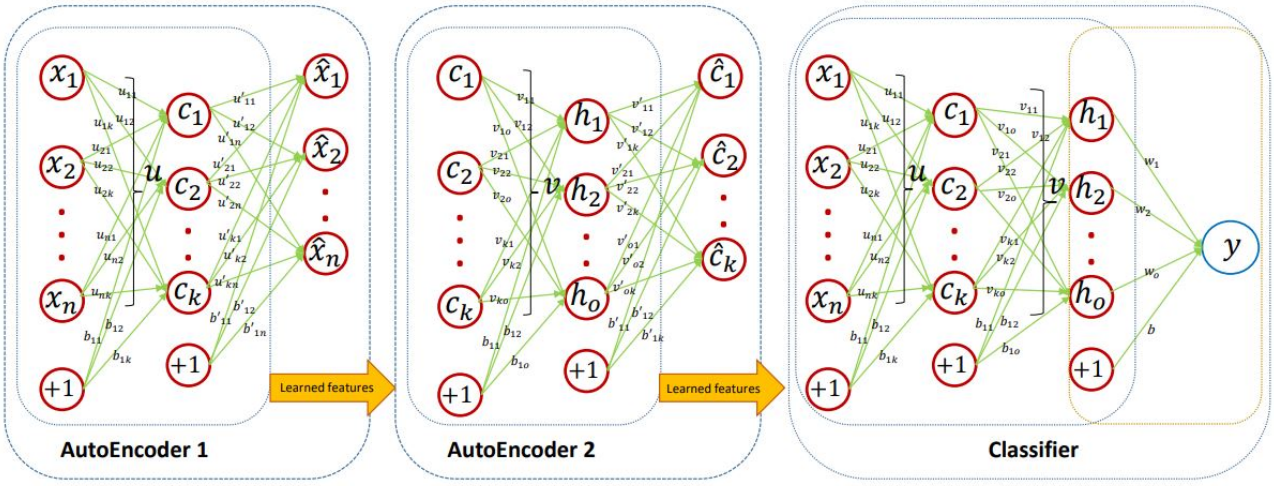

Figure 2. The architecture of proposed SAE-DNN (Here we have, n = 126, k = 64, o = 32. The matrix W is representative of the classifier, which is a four-layer DNN and is not shown here)

show total accuracy of 98%, a precision of 97.2%, a recall of 99%, and F-score of 0.9809.

We compared the results of our approach with unsupervised and supervised approaches, which used the WADI dataset to train and evaluate their models. As illustrated in Table 1. using only 2% of labeled data, the results of our approach (SAE-DNN) outperform the supervised systems (SVM, KNN, BYS) which use the full dataset to train and test their algorithms.

Our approach outperforms the best-unsupervised system among these unsupervised approaches, which shows the efficiency and usability of our proposed approach in a practical environment. We should note that the more percentage of the labeled dataset we use, the lower the False Positive Rate (FPR) we achieve. Table 2 illustrates different values of False Positive Rates for different percentages of labeled data usage. As shown in this table, the FPR decreases as the percentage of labeled data used increases. We see that using only 2% of labeled data, FPR is lower than 1%, which is a suitable rate for practical use cases.

As we mentioned earlier, another major challenge would be that any IDS that is implemented for industrial systems should work in real-time since most of them are critical-to-life, unlike IT, where the software patch comes in a while. Therefore, any IDS would need to be a real-time one. This challenge necessitates real-time detection and real-time data processing. The computation time of each test sample for our approach is about 60µsec, which shows the effectiveness of our proposed SAE-DNN approach.

It is notable that due to the repeated process of water distribution through the data collection phase, we removed the time feature and then deleted all repeated rows of data in order to avoid any malfunction when detecting real anomalies. In addition, due to the imbalanced dataset issue that may decline the detection rate of the system, we used SMOTE oversampling technique to balance the ratio of abnormal samples to normal samples in the dataset. This technique helped us with a significantly improved detection rate. We also removed the normal data up to one hour after each attack because of the fluctuating behavior of the system in that period to get back to the normal state of the system. In our experiments, we trained the autoencoders, classifier, and fine-tuner for 15 epochs, 100 epochs, and 60 epochs, respectively.

Table 1. Comparison of the proposed approach with some supervised and unsupervised approaches

	Accuracy	Precision	Recall	F-score
MAD-GAN [2]	66%	7%	99.98%	0.13
FID-GAN [3]	84.2%	81%	99.98%	0.895
TGAN [4]	94.3%	92%	99.98%	0.9583
CNN-AE [5]	95.1%	94.1%	98.6%	0.963
Dual IF [6]	81%	76.5%	99.94%	0.867
SVM [9]	97.1%	–	–	–
KNN [9]	99.8%	–	–	–
BYS [9]	93.5%	–	–	–
USMD [7]	96.2%	94.64%	99.76%	0.9702
SAE-DNN (1%)	98%	97.2%	99%	0.9809
SAE-DNN (2%)	99.3%	99.07%	99.5%	0.9928
SAE-DNN (3%)	99.57%	99.44%	99.71%	0.9957
SAE-DNN (10%)	99.8%	99.91%	99.98%	0.9994

6 Conclusion and Future Work

This paper described the proposed Stacked AutoEncoder Deep Neural Network (SAE-DNN) intrusion detection system for multisensory data to leverage in Cyber-Physical systems. SAE-DNN is a data-driven detection system, which presents a semi-supervised

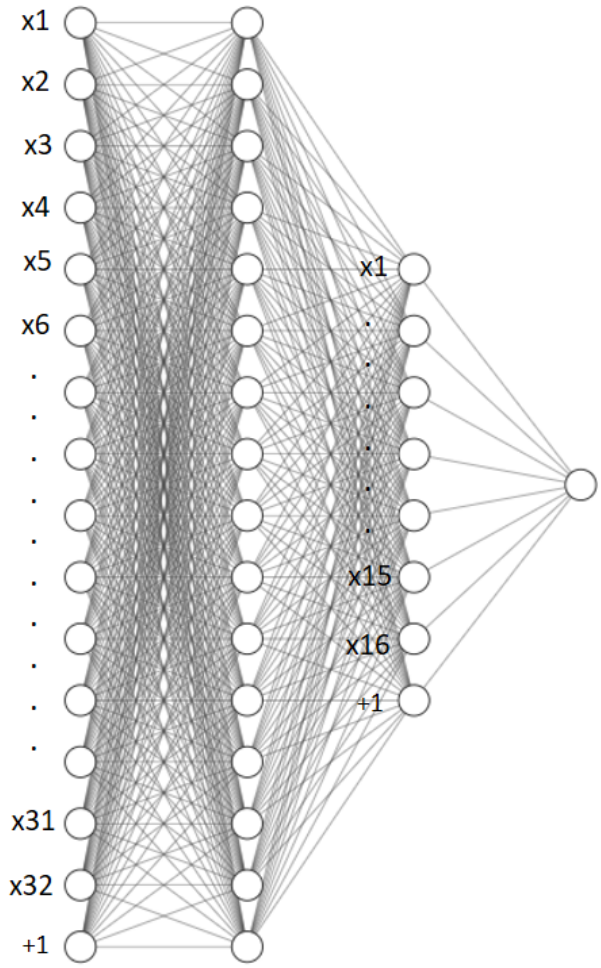

Figure 3. The architecture of the proposed DNN classifier

approach with good performance, and applies to a wide range of CPSs. SAE-DNN consists of a deep learning-based feature extractor and a deep neural network-based classifier, which can be exploited to detect anomalies with a high detection rate and a low false positive rate in a real-time process. The stacked autoencoder takes normal samples as input and trains itself to represent a low-dimension and low-noise representation of input data. Moreover, the proposed SAE-DNN uses a DNN classifier, which can be used to classify the anomalies and normal data with a high detection rate and low false positive rate using highly related features extracted from the deep neural network. In addition, the proposed approach includes a fine-tuning step, which helps us with a significantly improved detection rate of malicious data. We also showed the effectiveness of our proposed approach conducted on the WADI (a publicly available dataset collected from a real water distribution testbed) dataset using different detection metrics and compared it with the state-of-the-art methods. The results showed a good performance of SAE-DNN while using a few percentages of labeled data. The results of the proposed SAE-DNN using only 2% of the labeled dataset showed the superior performance of SAE-DNN over existing supervised and unsupervised methods. For future work, we intend to strengthen the security of our proposed system against adversarial attacks. As a future direction, promoting the security of our proposed scheme against adversarial attacks would be interesting. Furthermore, an attempt to sketch new roads to design unsupervised generative models might potentially result in more accurate IDSs. In addition, we intend to further study the design of unsupervised generative models, which helps us with data augmentation. This will help us to design new intrusion detection systems more accurately.

Table 2. Different values of False Positive Rates for different percentages of labeled data usage

	Accuracy	FPR
SAE-DNN (1%)	98%	2.85%
SAE-DNN (2%)	98%	0.92%
SAE-DNN (3%)	98%	0.83%
SAE-DNN (5%)	98%	0.68%
SAE-DNN (10%)	98%	0.24%

Acknowledgment

We appreciate the insightful comments from Mr. Mojtaba Shirinjani on editing the paper text.

References

[1] Sridhar Adepu, Venkata Reddy Palleti, Gyanendra Mishra, and Aditya Mathur. Investigation of cyber attacks on a water distribution system. In *Applied Cryptography and Network Security Workshops: ACNS 2020 Satellite Workshops, AIBlock, AIHWS, AIoTS, Cloud S&P, SCI, SecMT, and SiMLA, Rome, Italy, October 19–22, 2020, Proceedings 18*, pages 274–291. Springer, 2020.

[2] Dan Li, Dacheng Chen, Baihong Jin, Lei Shi, Jonathan Goh, and See-Kiong Ng. Mad-gan: Multivariate anomaly detection for time series data with generative adversarial networks. In *International conference on artificial neural networks*, pages 703–716. Springer, 2019.

[3] Paulo Freitas de Araujo-Filho, Georges Kaddoum, Divanilson R Campelo, Aline Gondim Santos, David Macêdo, and Cleber Zanchettin. Intrusion detection for cyber–physical systems using generative adversarial networks in fog environment. *IEEE Internet of Things Journal*, 8(8):6247–6256, 2020.

[4] Xueqin Zhang, Jiyuan Wang, and Shinan Zhu. Dual generative adversarial networks based unknown encryption ransomware attack detection. *IEEE Access*, 10:900–913, 2021.

[5] Hakan Kayan, Yasar Majib, Wael Alsafery, Mah-

moud Barhamgi, and Charith Perera. Anomliot: An end to end re-configurable multi-protocol anomaly detection pipeline for internet of things. *Internet of Things*, 16, 2021.

[6] Mariam Elnour, Nader Meskin, Khaled Khan, and Raj Jain. A dual-isolation-forests-based attack detection framework for industrial control systems. *IEEE Access*, 8:36639–36651, 2020.

[7] Abdullah Alsaedi, Zahir Tari, Redowan Mahmud, Nour Moustafa, Abdun Mahmood, and Adnan Anwar. Usmd: Unsupervised misbehaviour detection for multi-sensor data. *IEEE Transactions on Dependable and Secure Computing*, 20(1):724–739, 2022.

[8] Muhammad Omer Shahid, Chuadhry Mujeeb Ahmed, Venkata Reddy Palleti, and Jianying Zhou. Curse of system complexity and virtue of operational invariants: machine learning based system modeling and attack detection in cps. In *2022 IEEE Conference on Dependable and Secure Computing (DSC)*, pages 1–8. IEEE, 2022.

[9] Paola Perrone, Francesco Flammini, and Roberto Setola. Machine learning for threat recognition in critical cyber-physical systems. In *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, pages 298–303. IEEE, 2021.

[10] Tuan A Tang, Lotfi Mhamdi, Des McLernon, Syed Ali Raza Zaidi, and Mounir Ghogho. Deep learning approach for network intrusion detection in software defined networking. In *2016 international conference on wireless networks and mobile communications (WINCOM)*, pages 258–263. IEEE, 2016.

[11] Chuadhry Mujeeb Ahmed, Venkata Reddy Palleti, and Aditya P Mathur. Wadi: a water distribution testbed for research in the design of secure cyber physical systems. In *Proceedings of the 3rd international workshop on cyber-physical systems for smart water networks*, pages 25–28, 2017.

[12] S. Flores, "Variational Autoencoders are Beautiful," 15 April 2019. [Online]. Available: https://www.compthree.com/blog/autoencoder/.

Amirhosein Salehi received the B.Sc. degree in electrical engineering from Amirkabir University of Technology. His educational orientation was in the Bachelor's course in Telecommunication, and in the Master's course in Secure Telecommunication and Cryptography.

Siavash Ahmadi has completed all his studies at the Faculty of Electrical Engineering of Sharif University of Technology. his educational orientation was in the Bachelor's course in Telecommunication, in the Master's course in Secure Telecommunication and Cryptography, and in the Ph.D. course in Telecommunication System.

Mohamad Reza Aref received the M.Sc. and Ph.D. degrees in electrical engineering from Stanford University. He was a faculty member of the Isfahan University of Technology from 1982 to 1995. Since 1995, he has been a professor of electrical engineering with the Sharif University of Technology. His research interests include Communication and Information Theory, and Cryptography.