Cyber Security Intrusion Detection for Agriculture 4.0: Machine Learning-Based Solutions, Datasets, and Future Directions

Mohamed Amine Ferrag, Lei Shu, *Senior Member, IEEE*, Othmane Friha, and Xing Yang

***Abstract*—In this paper, we review and analyze intrusion detection systems for Agriculture 4.0 cyber security. Specifically, we present cyber security threats and evaluation metrics used in the performance evaluation of an intrusion detection system for Agriculture 4.0. Then, we evaluate intrusion detection systems according to emerging technologies, including, Cloud computing, Fog/Edge computing, Network virtualization, Autonomous tractors, Drones, Internet of Things, Industrial agriculture, and Smart Grids. Based on the machine learning technique used, we provide a comprehensive classification of intrusion detection systems in each emerging technology. Furthermore, we present public datasets, and the implementation frameworks applied in the performance evaluation of intrusion detection systems for Agriculture 4.0. Finally, we outline challenges and future research directions in cyber security intrusion detection for Agriculture 4.0.**

***Index Terms*—Agriculture 4.0, cyber security, intrusion detection system, machine learning approaches, smart agriculture.**

I. Introduction

THE agricultural and industrial revolution has evolved through the following four generations: Agriculture 1.0, Agriculture 2.0, Agriculture 3.0, and Agriculture 4.0, as depicted in Fig. 1. Agriculture 1.0 refers to the practices of agriculture from the beginning of human civilization until the end of the 19th century, a period when farmers depended heavily on traditional cultivation tools such as the traditional plough for creating favorable conditions for seed placement and plant growth. At the beginning of the 20th century, the increase in agricultural production was known as Agriculture 2.0 based on the agricultural machinery includes using combines, irrigation, harvesting, trucks, tractors, aircraft, helicopters, etc. From the early seventies to the present day, Agriculture 3.0 appeared which is based on green renewable energy such as bioenergy, geothermal energy, solar energy, hydropower, and wind power [1].

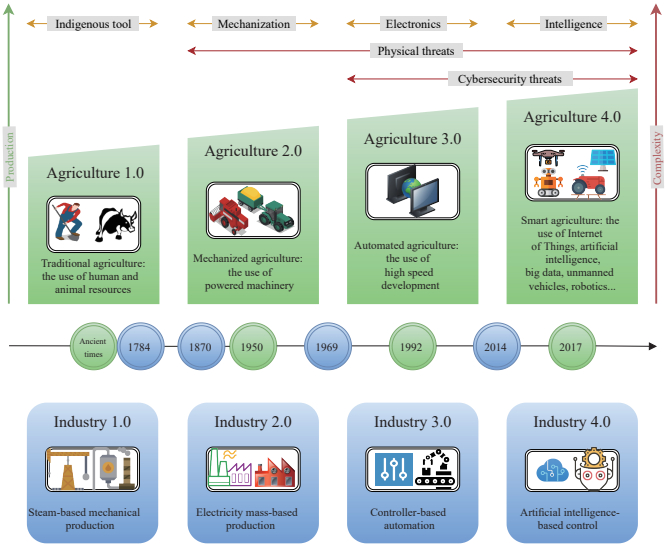

Fig. 1. The development of agricultural revolutions with industrial revolutions and related cyber security threats.

The term "Agriculture 4.0" appeared following "Industry 4.0" [2], [3], which is defined by a combination of technologies that are emerging such as Blockchain, software-defined networking (SDN), Artificial Intelligence, Internet of Things (IoT), IoT devices, 5G communications, Drones, Fog/Edge computing, Cloud computing, network function virtualization (NFV), Smart Grids, etc. The diagram of Agriculture 4.0 is shown in Fig. 2. In the physical layer, various IoT devices (e.g., sensor and camera) and drones are applied to monitor agricultural environmental conditions by collecting meteorological data, soil moisture, crop image, and livestock behavior analysis, and health monitoring data. Different actuators (e.g., autonomous tractors, insecticidal lamps, feeding machine, and irrigation equipment) are activated when the data meets specific conditions, which promote the automation of agriculture production and management. Besides, new energy technology (e.g., solar

Manuscript received March 17, 2021; revised April 19, 2021; accepted May 17, 2021. This work was supported in part by the Research Start-Up Fund for Talent Researcher of Nanjing Agricultural University (77H0603) and in part by the National Natural Science Foundation of China (62072248). Recommended by Associate Editor MengChu Zhou. *(Corresponding author: Lei Shu.)*

Citation: M. A. Ferrag, L. Shu, O. Friha, and X. Yang, "Cyber security intrusion detection for agriculture 4.0: Machine learning-based solutions, datasets, and future directions," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 3, pp. 407–436, Mar. 2022.

M. A. Ferrag is with Department of Computer Science, Guelma University, B.P. 401, 24000, Algeria (e-mail: ferrag.mohamedamine@univ-guelma.dz).

L. Shu is with the College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China, and also with the School of Engineering, University of Lincoln, Lincoln LN67TS, UK (e-mail: lei.shu@ieee.org).

O. Friha is with the Networks and Systems Laboratory (LRS), University of Badji Mokhtar-Annaba, B.P.12, Annaba 23000, Algeria (e-mail: othmane.friha@univ-annaba.org).

X. Yang is with the College of Engineering, Nanjing Agricultural University, Nanjing 210031, China (e-mail: harryyangx@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2021.1004344

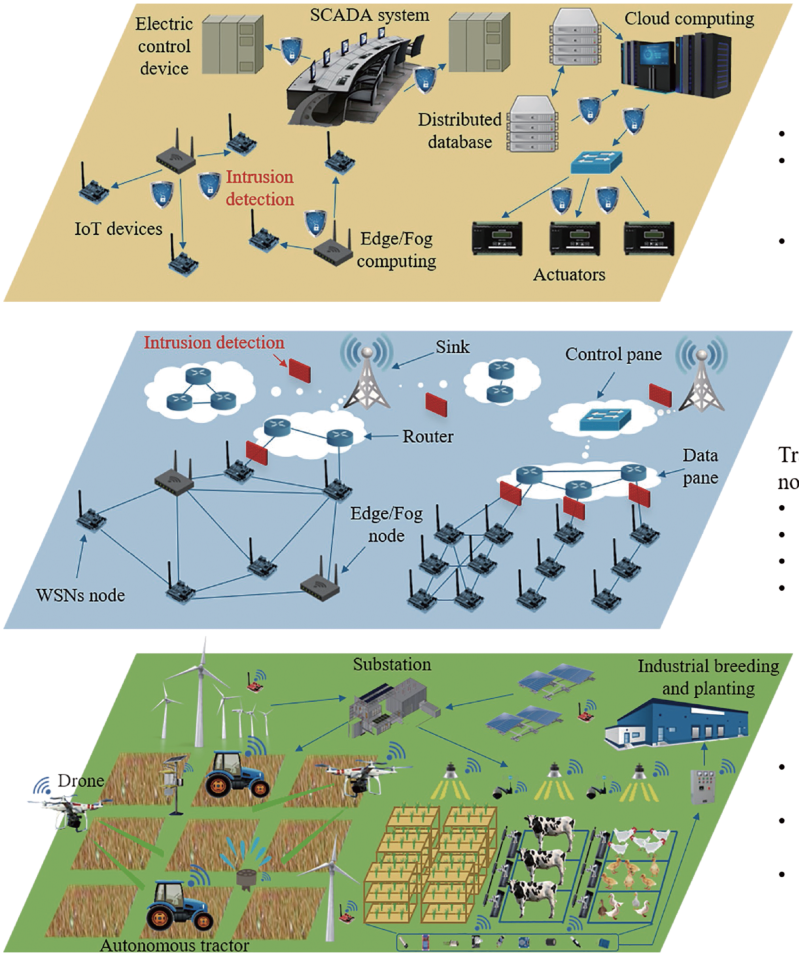

Fig. 2. The diagram of Agriculture 4.0, including 1) Cloud computing-enabled Agriculture 4.0, Edge/Fog computing-enabled Agriculture 4.0 in the application layer, 2) SDN/NFV network-enabled Agriculture 4.0 in the network layer, and 3) Smart Grid-enabled Agriculture 4.0, autonomous tractor-enabled Agriculture 4.0, IoT-device enabled Agriculture 4.0, Drone-enabled Agriculture 4.0, industrial Agriculture 4.0 in the physical layer.

power and wind power) and smart grid architecture contribute to supplying energy for IoT devices in Agriculture 4.0 [4]. In the network layer, intelligent agriculture devices transmit data to the Edge/Fog node and sink node by the wireless network, which forms different types of networks, e.g., the mesh network (ZigBee based), star network (LoRa based), GSM network (4G/5G based), and SDN (including control pane and data pane) [5]. In the application layer, cloud computing is applied in the decision-making of agriculture production and management by analyzing data, which are stored in a distributed database. Edge/Fog computing is used to implement tasks with higher timeliness requirements. Moreover, the supervisory control and data acquisition (SCADA) system is usually used to monitor the operation status of electrical equipment in a Smart Grid.

These emerging technologies have been widely applied in Industry 4.0, and it is not difficult to imitate application in agricultural scenarios. Therefore, the major challenge of developing Agriculture 4.0 does not reside in the deployment of the emerging technologies, but primarily in the guarantee of security and privacy, since the deployment of thousands of IoT-based devices is in an open field. In addition, there are many security and privacy issues associated with each layer of the IoT architecture [6]. For example, an adversary can initiate many cyberattacks, such as distributed denial-of-service (DDoS) attacks to make a service unavailable and then inject false data, which affects food safety, agri-food supply chain efficiency, and agricultural productivity. The research community in cyber security proposes the use of intrusion detection systems (IDS), which is a technology for the security of networks that are dedicated to continuously observing events inside a computing or networking system, and then evaluate them against intrusion evidence [7], [8]. To further protect Agriculture 4.0 from cyber attacks, the IDS can be implemented in conjunction with other security solutions including, encryption techniques, authentication, authorization, and Blockchain [9].

To detect malicious behaviors, the IDSs use artificial intelligence-based techniques, such as hybrid machine learning, voting based extreme learning machine, deep learning techniques, hierarchical approaches, reinforcement learning, etc. The IDSs based on machine learning techniques have been covered by many surveys. Table I presents the related surveys on the IDSs based on machine learning techniques. Many surveys focused on IDSs based on deep learning approaches [11], [12], [16], [19], machine learning techniques [10], [13], and support vector machines [20]. Some surveys payed attention to SCADA systems [21], SDN

TABLE I
RELATED SURVEYS ON THE IDSS BASED ON MACHINE LEARNING TECHNIQUES

Year	Authors	Taxonomy	IDS building process for Agriculture 4.0	Public datasets	Benefits of IDS for Agriculture 4.0	Open challenges and future research opportunities for Agriculture 4.0
2016	Buczak and Guven [10]	- Machine learning techniques	No	Partial	No	No
2019	Kwon *et al*. [11]	- Deep learning techniques	No	No	No	No
2020	Al-Garadi *et al*. [12]	- Deep learning techniques	No	No	No	No
2019	Mishra *et al*. [13]	- Machine learning techniques	No	Yes	No	No
2019	da Costa *et al*. [14]	- Machine learning techniques	No	Partial	No	No
2019	Chaabouni *et al*. [15]	- IoT threats classification	No	Yes	No	No
2019	Liu and Lang [16]	- Machine learning and deep learning techniques	No	No	No	No
2019	Sultana *et al*. [17]	- SDN	No	No	No	No
2020	Ahmad *et al*. [18]	- SDN	No	No	No	No
2020	Ferrag *et al*. [9]	- Deep learning techniques	No	Yes	No	No
2021	Ahmad *et al*. [19]	- Machine learning and deep learning techniques	No	No	No	No
2021	Mohammadi *et al*. [20]	- Support vector machines	No	No	No	No
	Our survey	- Cloud computing-enabled Agriculture 4.0 - Fog/Edge-enabled Agriculture 4.0 - SDN/NFV-enabled Agriculture 4.0 - Drones-enabled Agriculture 4.0 - Autonomous tractors-enabled Agriculture 4.0 - IoT devices-enabled Agriculture 4.0 - Industrial Agriculture 4.0 - Smart Grid-enabled Agriculture 4.0	Yes	Yes	Yes	Yes

TABLE II
RELATED SURVEYS ON AGRICULTURE 4.0

Year	Authors	Public datasets	Intrusion detection systems	Machine learning and deep learning approaches	Main focus/contributions
2017	Ray [22]	No	No	No	IoT deployments in terms of hardware platforms and communication technologies
2018	Kamilaris and Prenafeta-Boldú [23]	No	No	Yes	A review on the deep learning approaches applied in agriculture
2018	Elijah *et al*. [24]	No	No	No	An overview of data analytics and IoT in agriculture
2019	Khanna and Kaur [25]	No	No	No	A review of IoT in the field of precision agriculture
2020	Zhai *et al*. [26]	No	No	No	Feasibility of decision support systems for Agriculture 4.0
2021	Liu *et al*. [1]	No	No	No	Address the main applications of evolving technologies in the agricultural sector such as big data analytics, robotics, Artificial Intelligence, etc.
2021	Yang *et al*. [27]	No	No	No	Discuss security and privacy challenges as well as technologies and development modes in Smart Agriculture
2021	Friha *et al*. [28]	No	No	No	Review emerging technologies for IoT-based Intelligent Agriculture.
	Our Survey	Yes	Yes	Yes	A survey that covers IDS models, public datasets, and deep learning approaches

technology [17], [18], and IoT networks [14], [15]. In contrast, this survey proposes seven taxonomies that are related to 1) Cloud computing-enabled Agriculture 4.0, 2) Fog/Edge-enabled Agriculture 4.0, 3) SDN/NFV-enabled Agriculture 4.0, 4) Drone-enabled Agriculture 4.0, 5) Autonomous tractor-enabled Agriculture 4.0, 6) IoT device-enabled Agriculture 4.0, 7) Industrial Agriculture 4.0, and 8) Smart Grid-enabled Agriculture 4.0. It also provides a more comprehensive review by covering novel security topics such as the IDS building process, public datasets, benefits of IDS, and open challenges and future research opportunities for Agriculture 4.0. Therefore, some surveys cover different aspects of Agriculture 4.0. As shown in Table II, we classify the related studies by the following factors:

1) IDSs: It indicates whether the survey provided a taxonomy of IDSs for Agriculture 4.0.

2) Public Datasets: It specifies if the survey presented the public datasets used in the performance evaluation of IDSs for Agriculture 4.0.

3) Machine Learning and Deep Learning Techniques: It

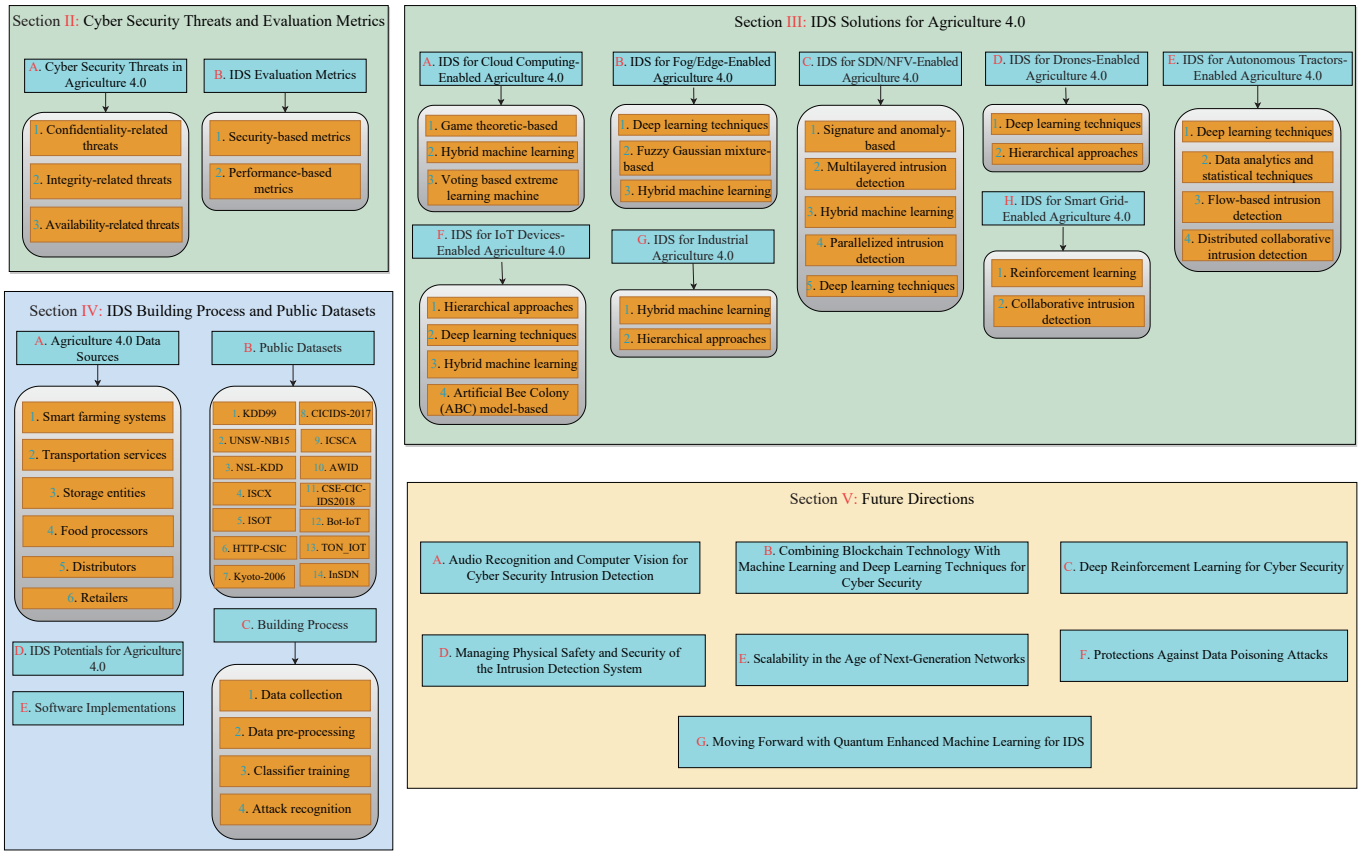

Fig. 3. Organization of the content in the rest of this article.

specifies if the survey has provided a comparative study and evaluated machine learning and deep learning approaches for cyber security intrusion detection in Agriculture 4.0.

Most of the surveys on Agriculture 4.0 outline the new technologies [1], [22]–[26], [28] such as Fog/Edge computing, software-defined networking, network function virtualization, and unmanned aerial vehicles, without focusing on the performance of machine learning and deep learning techniques for cyber security. To the best of our knowledge, our survey is the first that thoroughly covers the performance of machine learning techniques for cyber security in Agriculture 4.0.

Our survey differs from the earlier-mentioned works in the following points:

1) We present cyber security threats and evaluation metrics used in the performance evaluation of IDSs for Agriculture 4.0.

2) We provide a comprehensive classification and in-depth analysis of machine learning and deep learning based IDSs for cyber security in Agriculture 4.0.

3) We provide a detailed description of the current best practices, implementation frameworks, and public datasets used in the performance evaluation of IDSs for Agriculture 4.0.

4) We highlight remaining challenges and future research directions in cyber security intrusion detection for Agriculture 4.0.

As shown in Fig. 3, the remainder of this paper is structured as follows. Section II focuses on Agriculture 4.0 cyber security threats and IDS evaluation metrics. Section III presents the IDS solutions for Agriculture 4.0. Section IV describes IDS building process and public datasets. Section V discuss the future directions. Lastly, Section VI presents the conclusions.

II. Cyber Security Threats and Evaluation Metrics

Although Agriculture 4.0 is envisioned to be the new standard, challenges to its acceptance and widespread adoption may be constrained by potential threats. Traditionally, some of those threats persist throughout the years, such as rough weather conditions. Although others are attributed to the broad development of technological solutions, resulting in major security gaps and serious attack vectors such as ransomware, supply chain attacks, IoT attacks, and many others [27], [29].

A. Cyber Security Threats in Agriculture 4.0

The U.S. Department of Homeland Security defined three major cyber threat categories for Precision Agriculture, namely: confidentiality-related, integrity-related, and availability-related threats [30].

1) Confidentiality-Related Threats: With a variety of communication devices in intelligent agriculture, data flows through multiple interconnected devices from source to destination. Privacy threats can lead to loss of privacy and violations of data or information [29]. As farmers are highly protective of their information, for example, data on yields, farmland prices, and livestock health, it is important that this

information be kept confidential. Losing or misusing this data has potentially dramatic financial, emotional and reputational consequences on farmers [30].

2) Integrity-Related Threats: Collecting and using data is an important step in helping farmers make real-time intelligent management decisions. Due to potential unauthorized or inappropriate alterations to the reliability of data or resources, it is possible that information from Intelligent Agricultural Systems can become unreliable or inaccurate, and may result in possible financial abuse [29].

3) Availability-Related Threats: Failing to provide available services to customers can cause business disturbances, and possible loss of customer confidence and earnings. For example, if an attacker stopped the activities of an existing Intelligent Agriculture Network food security would be impacted and there would be a serious loss of reputation for the equipment manufacturer [30].

B. IDS Evaluation Metrics

Several metrics can be used to evaluate the efficiency and effectiveness of an IDS, and most of them fall into one of the following categories: security-based metrics and performance-based metrics [31], [32].

1) Security-Based Metrics: Metrics in this category describe the effectiveness of the IDS to determine the distinction between intrusive and non-intrusive activities. Being a binary classifier, an IDS can result in one of the following outputs: *a) true positive (TP):* when an intrusion is correctly classified as an intrusive action; *b) true negative (TN):* when a legitimate action is properly classified as legitimate; *c) false positive (FP):* when a legitimate action is erroneously classified as an intrusion; and *d) false negative (FN):* when an intrusion is erroneously classified as a legitimate action [31]. Well-known metrics within this category include:

- *Confusion matrix:* This metric reflects the result of the classification. For instance, it represents the true and false results of the classification. It can have 2×2 dimensions in the case of binary classification, but it can also have $N\times N$ dimensions in the case of a multi-class classifier with N different classes. Although the confounding matrix is not a metric alone, rather it is a baseline of metrics from which other indicators of effectiveness can be quantified.
- *Accuracy:* This metric is essentially the correct classification rate of an IDS, whether for validation set or test set. Accuracy is obtained with

$$\frac{TP+TF}{TP+TF+FP+FN}. \tag{1}$$

- *Precision:* This metric represents the ratio of the classified actions by the IDS that are intrusive. Precision is obtained with

$$\frac{TP}{TP+FP}. \tag{2}$$

- *Recall:* This metric is the ratio of intrusive actions classified by the IDS as intrusive. The recall is obtained with

$$\frac{TP}{TP+FN}. \tag{3}$$

- *F_β-score:* This metric is a weighted harmonic mean of precision and recall, where β mirrors the significance of the recall concerning accuracy. The F-score is also applied when evaluating a multi-class classification. The F_1-score is obtained with (4). The final F_1-score is obtained by micro-averaging based on class frequency (mico-F_1), or by macro-averaging based on the same importance of all classes (macro-F_1) [33]. The F_1-score is frequently compared to the G-measure which is the geometric mean of the precision and recall, and it is used to evaluate binary and multi-class classifiers in cases of class imbalance [34].

$$2\times\frac{Precision\times Recall}{Precision+Recall}. \tag{4}$$

- *ROC curve:* The receiver operator characteristic (ROC) curve is a robust metric that shows the sensitivity and specificity associated with a continuous variable. it is a plot of coordinates composed of true positive rate (TPR) (3), a vertical axis, and false positive rate (FPR) (5), a horizontal axis. The area under the ROC curve, called AUC, is regarded as a key evaluation measure.

$$\frac{FP}{FP+TN}. \tag{5}$$

2) Performance-Based Metrics:

i) Computational cost: the computational cost represents the amount of time required to accomplish an essential task to classify an action as intrusive or legitimate.

ii) Communication overhead: is the volume of data that can be processed by an IDS per second. This means the throughput rate expressed in Giga Bits per second to assert the performance displayed by the IDS.

iii) CPU usage: this metric represents the overhead rate on the CPU when adding an IDS to the infrastructure.

iv) Memory usage: this metric represents the memory consumption required by an IDS to do its classification.

v) Energy consumption: this metric represents the extra energy consumed by a device when an IDS is introduced. This measure is essential for hardware-limited appliances such as mobile and IoT devices.

III. IDS Solutions for Agriculture 4.0

Agriculture 4.0 uses many emerging technologies such as SDN/NFV, Cloud computing, Fog/Edge computing, Drones, Autonomous tractors, IoT devices, Smart Grids, etc. According to these emerging technologies, we review and analyze the IDSs that use machine learning and deep learning techniques for cyber security in Agriculture 4.0.

A. IDS for Cloud Computing-Enabled Agriculture 4.0

Using various IoT devices in Agriculture 4.0, there are many new threats in the cloud environment as these devices are very susceptible to security attacks. The IDSs for Cloud computing-enabled Agriculture 4.0 can be categorized into three classes, namely, 1) Game theory-based, 2) Hybrid machine learning, and 3) Voting based extreme learning machine.

1) Game Theory-Based: Although game theory is not a branch of machine learning, this technique has shown very

good results in IDS. Gill *et al*. [35] developed a model of game theory, named GTM-CSec, to provide intelligent detection of attacks in the cloud environment. The GTM-CSec model is composed of two main components: cooperative and non-cooperative games and is based on three techniques, namely, the signature, anomaly, and honeypot techniques. These techniques are adopted in the following four components: perception, logical analysis, computational analysis, and decisive analysis. The performance evaluation in MATLAB with payoff functions and probabilities showed that the GTM-CSec model can improve electricity usage of the defense mechanism and is very efficient in protecting against attackers.

2) Hybrid Machine Learning: Based on monitoring user patterns of behavior, Rabbani *et al*. [36] designed a hybrid machine learning system, which is based on extracting users' behavioral patterns to detect malicious behaviors in the cloud computing environment. To construct a network that is automatically optimized, the proposed system uses particle swarm optimization-based probabilistic neural networks (PSO-PNN). The study used UNSW-NB15 dataset where the features are presented in a quantitative (i.e., numerical) and qualitative (i.e., symbolic) format. The experimental results reported that the PSO-PNN approach provides high accuracy to detect suspicious activities.

3) Voting Based Extreme Learning Machine: Kushwah and Ranga [37] considered a cloud infrastructure with a detector attached, which is based on three components, namely, i) training database component, ii) preprocessor component, and iii) classifier component. The detector uses a voting extreme learning machine to identify DDoS threats in a cloud computing framework, as presented in Fig. 4. The study used two datasets, namely, i) the NSL-KDD dataset and ii) the ISCX dataset. The experimental results reported that the proposed system provides high accuracies of 99.18% and 92.11% with the NSL-KDD and ISCX datasets, respectively. Aldribi e*t al*. [38] designed an IDS based on a hypervisor using online multivariate statistical change analysis for detecting anomalous cloud behavior. The study used the ISOT-CID dataset to validate cloud intrusion detection framework proposed. The experimental results reported that the proposed system's overall detection rate was 96.23% and false-positive rate of 7.56%.

B. IDS for Fog/Edge-Enabled Agriculture 4.0

Rather than transferring data generated by IoT-connected devices to the Cloud or a Data Center, Fog/Edge Computing in Agriculture 4.0 involves processing data at the edge of the network directly where it is generated [39]. The IDSs for Fog/Edge-enabled Agriculture 4.0 can be categorized into three classes, namely, 1) Deep learning techniques, 2) Fuzzy Gaussian mixture-based, and 3) Hybrid machine learning.

1) Deep Learning Techniques: To secure multiple web applications in fog computing, Tian et *al*. [40] proposed a distributed deep learning system using convolutional neural networks, which is a deep learning technique. Specifically, the system takes advantage of analyzing URLs, which can differentiate normal queries from those that are anomalous.

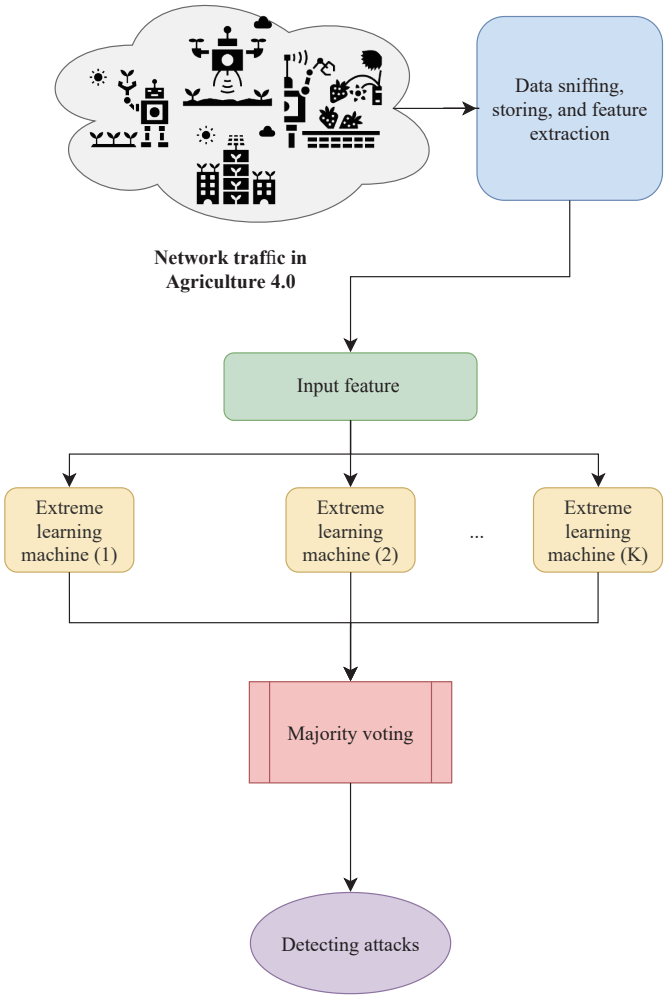

Fig. 4. Voting and extreme learning machine-based IDS for Agriculture 4.0.

The study used three datasets, namely, HTTP Dataset CSIC 2010, FWAF, and HttpParams Dataset to validate the proposed edge IDS. The experimental results reported that the accuracy fluctuates approximately 0.955. Almogren [41] proposed a malicious activity detection model for edge-of-things computing, which is based on deep belief networks. The proposed system followed three steps, namely, i) network data collection, ii) feature extraction, and iii) classification. The network data collection consists of collecting data and dividing it into training data and test data. The feature extraction consists of extracting features related to intrusion, while the classification consists of using these features to train a deep belief network. The study used the UNSW-NB15 dataset to validate the proposed edge intrusion detection framework, in which the experimental results reported that a deep belief network has the best overall performance as compared to support vector machine and artificial neural network. Jiang *et al*. [42] designed a system to detect and prevent identity theft attacks, named PHYAlert, for wireless edge networks. The PHYAlert system used the authenticity of the 802.11 data frame for spoofing detection (i.e., man-in-the-middle attack).

Wu *et al*. [43] developed a system for intrusion detection,

called SRDLM, which is based on semantic re-encoding and deep learning. The SRDLM system can be applied for Fog/Edge-based Agriculture 4.0 using three steps, namely, network traffic semantic re-encoding, a deep learning model, and multi-space projection. Ahsan *et al*. [44] designed an IDS based on robust adaptive multivariate hotelling's control chart. The proposed system uses two main steps, namely, i) the data preparation step and ii) the construction of a control chart. The data preparation step consists of Data acquisition and searching. The building of a monitoring system is separated into two stages: i) Phase I: Building a Normal Profile and ii) Phase II: Detection. The KDD99, NSL-KDD, and UNSW-NB 15 datasets are used in performance evaluation. The findings indicate that the proposed chart can reach an accuracy of about 98% for the KDD99 dataset. Qureshi *et al*. [45] developed a system for intrusion detection based on a deep sparse auto-encoder and self-taught learning and used the NSL-KDD dataset containing 41 features and attacks categorized into the following four attack categories: Probing attacks, Remote-to-local attacks, User-to-root attack, and Denial-of-service attack. The study demonstrated that the sparse automatic encoder trained on the enhanced feature space is more robust and stable than the one trained on the original feature space. Furthermore, the study results indicated that the proposed IDS is robust and provides improved prediction accuracy.

2) Fuzzy Gaussian Mixture-Based: Despite the fact that fuzzy Gaussian mixture is not a class of machine learning, the technique has proven to be very successful in IDS. To identify zero-day attacks in Fog/Edge-based Agriculture 4.0, the FGMC-HADS method proposed by Haider *et al*. [46] can be applied, which are fuzzy Gaussian mixture-based correntropy models. The FGMC-HADS method uses the following steps: i) Learning the raw data from the operating system's kernel, ii) Generating representative sequences by using the joint feature construction module, and iii) Classifying the sequences as normal or abnormal using the Gaussian mixture model. To determine the performance of the FGMC-HADS method, three Linux host datasets are used, namely, the Linux dataset of ToN_IoT [47], KDD-98 [48], and NGIDS-DS [49]. The results demonstrated the superiority of the FGMC-HADS method in terms of accuracy and error compared to machine-learning approaches such as the support vector machine and k-nearest neighbor. Naik *et al*. [50] designed an intrusion identification method, named E-TLBO-FLANN, which is based on a meta-heuristic and functional link neural network and can be applied for Fog/Edge-based Agriculture 4.0. The proposed E-TLBO-FLANN method incorporates the concept of elitism for enhancing the model's outcome. The KDDCup99 dataset is used to assess performance on detecting intrusive behavior, in which the results showed that the proposed E-TLBO-FLANN method is better than other competing techniques such as Toosi and Kahani [51] and Pfahringer [52].

The adversarial attacks can be a threat if it is used in Fog/Edge-based Agriculture 4.0, for instance, to change the classification of agricultural products or water quality. The IDS model proposed by Pawlicki *et al*. [53] can be applied to handling adversarial attacks against artificial neural networks. The study used an artificial neural network with 3 hidden layers and 40 neurons on the first hidden layer, 40 on the second hidden layer, and 20 on the third hidden layer. The results showed that a random force achieves results with higher recall and better precision compared to artificial neural networks. For detecting unknown web attacks using a hybrid IDS, Kaur and Singh [54] designed a deep learning-based system, named D-Sign, which is based on deep recurrent neural networks. The D-Sign system is based on three parts, namely, i) signature generation engine, ii) anomaly detection engine, and iii) misuse detection engine. Both NSL-KDD and CICIDS 2017 datasets are used in evaluating the performance, in which the results demonstrated that the D-Sign system achieves 99.1% and 99.4% instances correctly for CICIDS and NSL-KDD datasets, respectively.

3) Hybrid Machine Learning: The hybrid machine learning-based IDS for Agriculture 4.0. is presented in Fig. 5. Hosseini and Zade [55] proposed a hybrid IDS, named MGA-SVM-HGS-PSO-ANN, that can be applied for Fog/Edge-based Agriculture 4.0. The MGA-SVM-HGS-PSO-ANN system is based on two parts, a part to select features and a part to detect attacks. The feature selection part consists of combining features of a genetic algorithm and support vector machine with multi-parent crossover and multi-parent mutation. The attack detection part uses an artificial neural network combined with a particle swarm optimization and a hybrid gravitational search. The performance evaluation on the NSL-KDD dataset showed that the MGA-SVM-HGS-PSO-ANN system achieves a detection accuracy of 99.3%. Teng *et al.* [56] proposed a 2-class SVM and decision trees based collaborative and adaptive IDS, which can be applied in the cloud (global agent with a big group) and edge (local agent with a small group).

C. IDS for SDN/NFV-Enabled Agriculture 4.0

Both SDN and NFV technologies in Agriculture 4.0 offer a software-based management platform for controlling standard network hardware such as Big Switch Networks or Vmware NSX. With the diversity of network attacks against Agriculture 4.0, the SDN/NFV-based Agriculture 4.0 faces many security issues. The IDS for SDN/NFV-enabled Agriculture 4.0 can be categorized into five classes, namely, 1) Signature and anomaly-based, 2) Multilayered intrusion detection, 3) Hybrid machine learning, 4) Parallelized intrusion detection, and 5) Deep learning techniques.

1) Signature and Anomaly-Based: Although anomaly-based detection has surpassed signature-based detection, in some situations their integration has been very successful in IDSs. Ngo *et al*. [57] conceptualized an SDN-based architecture for secure forwarding devices using signature and anomaly-based IDSs, which can be applied for Agriculture 4.0, as presented in Fig. 6. Specifically, the proposed architecture integrated two intrusion detection engines, called F-NIDS and F-ANIDS. The F-NIDS engine used snort rules for classifying attack packets, while the F-ANIDS engine used machine learning algorithms. The performance evaluation of both intrusion detection engines is implemented on parallel hardware

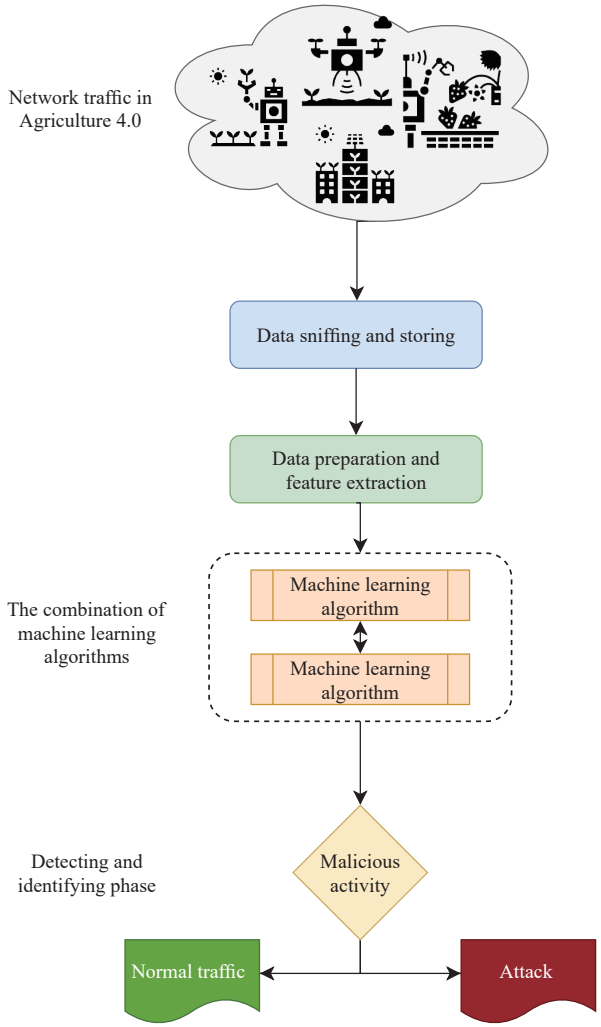

Fig. 5. Hybrid machine learning-based IDS for Agriculture 4.0.

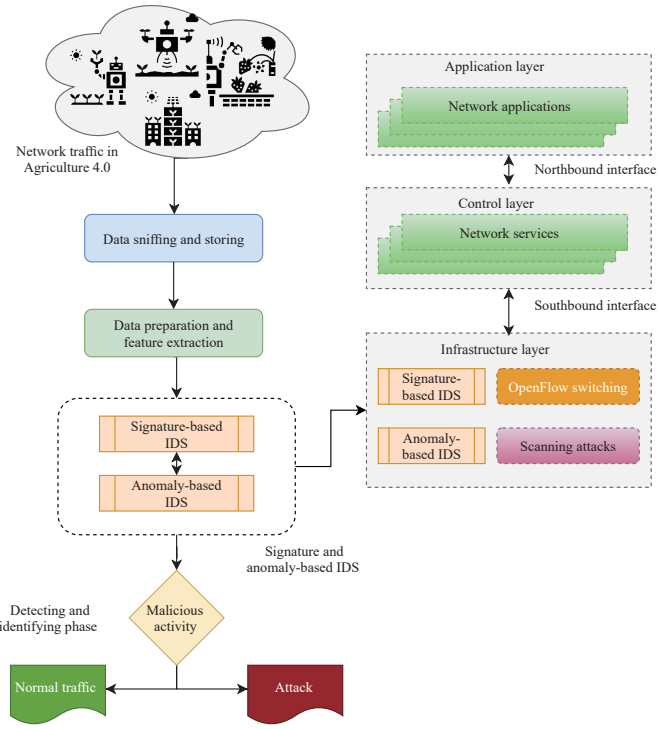

Fig. 6. Signature and anomaly-based IDS for SDN/NFV-enabled Agriculture 4.0.

platforms with two NetFPGA-10G boards (Xilinx xc5vtx240t FPGA device) and one GPU (GTX Geforce 1080 G1). The results showed that both intrusion detection engines are 14× faster using the GPU compared to using the CPU. It is important to note that signature-based detection has two major drawbacks [58], first, it is difficult to identify events that lead to an actual intrusion into network systems due to the massive amount of data logs, and second, the signature database grows exponentially over time. Ahmadon *et al*. [58] suggested a methodology based on Petri net-based [59] to detect intrusion behaviors, while reducing the number of alerts, along with an update method fusing two or more models of similar intrusion behaviors. The experiments performed showed the effectiveness of these methods.

2) Multilayered Intrusion Detection: Abdulqadder *et al*. [60] developed a multi-layered intrusion detection and prevention system, named MLP-IDP, for SDN/NFV enabled cloud of 5G networks. Specifically, the MLP-IDP proposed system is organized on five layers, namely, a switches layer, smart controller layer, domain controllers layer, data acquisition layer, and virtualization layer. In the data acquisition layer, the MLP-IDP system uses the Four-Q-Curve algorithm for authenticating mobile users with a trusted third party. To observe every switch in the data plan, the MLP-IDP system adopts an intrusion detection and prevention system which uses a deep reinforcement learning algorithm. In the domain controllers layer, the MLP-IDP system uses the Shannon Entropy function to classify the packets into normal or suspicious classes. In a smart controller layer, the MLP-IDP system uses multiple self-organizing maps to detects a DDoS attack. Compared to Abdulqadder *et al*.'s scheme [61], the MLP-IDP system is efficient in terms of security between switches and controllers. For visualizing network intrusion detection data, Zong *et al*. [62] designed an interactive method, which can improve the comprehension of network intrusion detection data in SDN/NFV-based Agriculture 4.0 through the use of a graphic display to show the relationship among the different categories of network traffic. Mishra *et al*. [63] provided a defensive mechanism against DDoS attempts which is based on entropy variations between a DDoS attack and regular traffic under low false positive rates and with slight processing overhead. All the simulations were performed inside a Mininet emulator using the POX controller. The proposed mechanism resulted in a detection rate of 98.2% and 0.04% false-positive rate.

3) Hybrid Machine Learning: Derhab *et al*. [64] proposed an intrusion detection framework, named RSL-KNN, which integrates the Blockchain and the software-defined network technologies. The RSL-KNN framework can be applied for SDN/NFV-based Agriculture 4.0. To defend against the forged command, the RSL-KNN framework uses two machine learning algorithms, namely, Random Subspace Learning and K-Nearest Neighbor. To prevent the misrouting attack in SDN/NFV-based Agriculture 4.0, the proposed framework

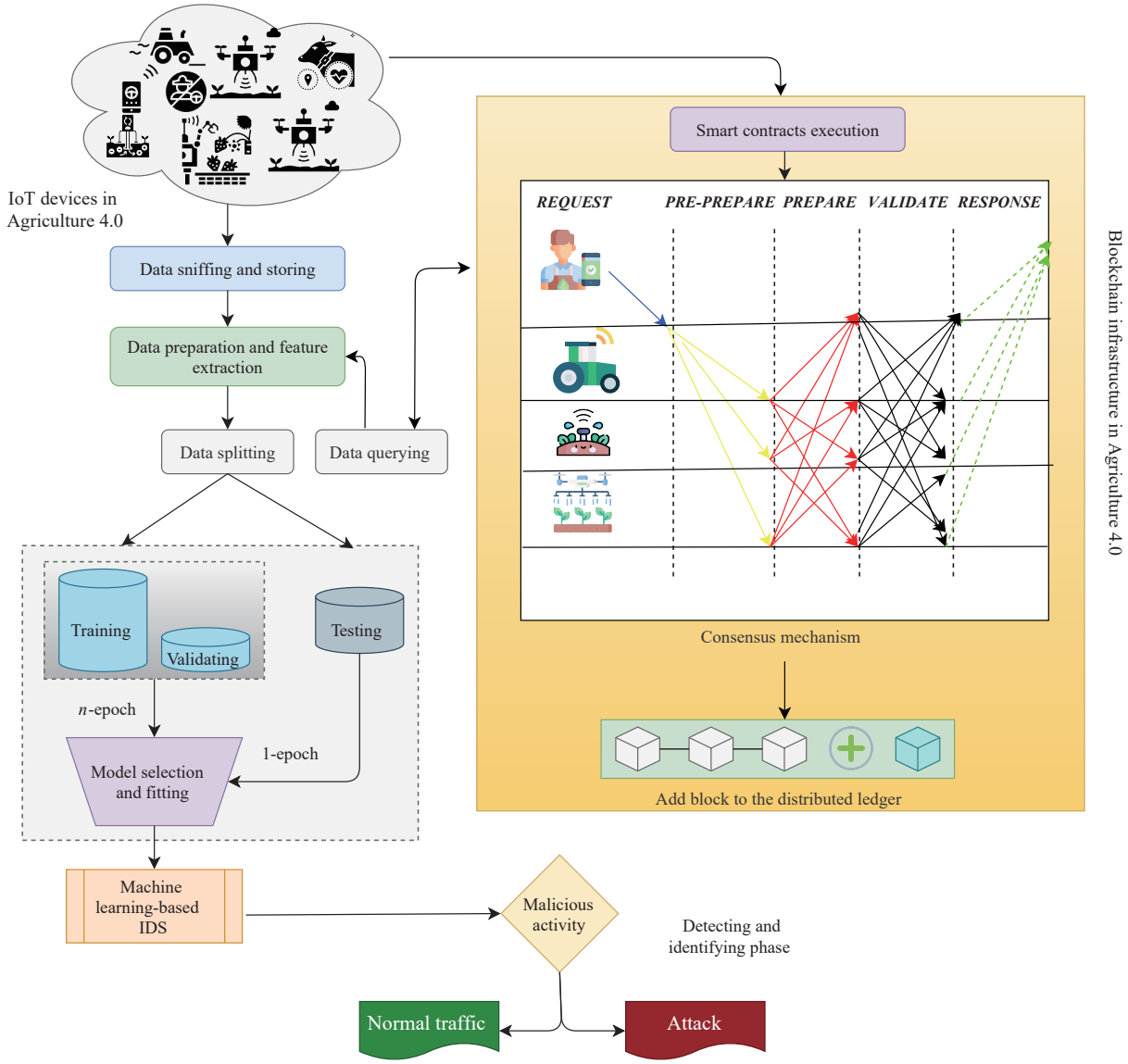

Fig. 7. Combination of blockchain and machine learning for cyber security intrusion detection for Agriculture 4.0.

uses a Blockchain-based integrity checking system. The performance evaluation on the Industrial Control System Cyber attack Dataset showed that the RSL-KNN framework system achieves detection accuracies of 91.07% and 96.73% under multi-class and binary class classification, respectively. The combination of blockchain and machine learning-based IDSs for Agriculture 4.0 cyber security is shown in Fig. 7. Based on feature selection and ensemble classifier techniques, Zhou *et al*. [65] introduced an IDS, which can be applied for SDN/NFV-based Agriculture 4.0. The heuristic algorithm is used for dimensionality reduction. Both C4.5 and Random Forest are used as classifier techniques for attack recognition. The NSL-KDD, AWID, and CIC-IDS2017 datasets are used in the experimental phase with Weka 3.8.3, where the results showed that the proposed system achieves detection accuracies of 98.3% and 99.3% for C4.5 and Random Forest classifiers, respectively. Lv *et al*. [66] used an extreme learning machine with a hybrid kernel function to build an intrusion detection approach, named KPCA-DEGSA-HKELM, which can be applied for SDN/NFV-based Agriculture 4.0. To detect attacks, the KPCA-DEGSA-HKELM system, a hybrid algorithm that combines the gravitational search algorithm and differential evolution algorithm, is used. The evaluation of performance using three datasets, namely, the industrial intrusion detection dataset, UNSW-NB15 dataset, and KDD99 dataset, shows that the KPCA-DEGSA-HKELM system can achieve higher computational efficiency with savings of 82.21%.

To improve the rate of precision in intrusion activities, Velliangiri and Karthikeyan [67] designed a hybrid optimization scheme based on adaptive particle swarm optimization and adaptive artificial bee colony optimization techniques. The hybrid optimization scheme uses four stages: i) Choice of the dataset, ii) Preprocessing of information, iii) Choice of feature, and iv) Hybrid categorization. The performance evaluation on the NSL-KDD dataset shows that the precision of the hybrid optimization scheme is increased to 94.23% and 97.85% compared to naive bayes and the support vector machine.

4) Parallelized Intrusion Detection: To ensure real-time ano-

maly detection, Chellammal and Malarchelvi [68] designed a parallelized intrusion detection architecture, which can be applied for SDN/NFV-based Agriculture 4.0. The proposed architecture is composed of five major components, namely, a model retraining component, prediction aggregator, ensemble-based prediction model, data partitioning component, and feature reducer. The performance evaluation on the three following datasets: KDD CUP 99, NSL - KDD, and Koyoto 2006 datasets, shows that the proposed architecture achieves an anomaly detection rate between 98% to 99%.

To improve the performance of classification and minimize calculation times, Khammassi and Krichen [69] proposed an NSGA2-LR wrapper approach, which can be applied for SDN/NFV-based Agriculture 4.0. The proposed wrapper approach is evaluated under two different frameworks: the first one employs a binomial logistic regression with numerous binary-class datasets for every attack type, and the second one employs a multinomial logistic regression with a multi-class dataset. The obtained results during the performance evaluation on the NSL-KDD dataset, UNSW-NB15 dataset, and CIC-IDS2017 dataset, showed a higher accuracy while using binary class datasets compared to multi-class datasets.

5) Deep Learning Techniques: The deep learning-based IDS for Agriculture 4.0. is presented in Fig. 8. Based on the combination of a conventional learning classifier system with a convolutional neural network, Bu and Cho [70] proposed a convolutional neural-based learning classifier system for IDSs, which can be applied for SDN/NFV-based Agriculture 4.0. To increase the detection rate of unfamiliar attacks, Yang *et al*. [71] designed a network intrusion detection model, called SAVAER-DNN, which can detect unknown and known attacks. The SAVAER-DNN model integrates the supervised variational auto-encoder data generation (SAVAER) and the wasserstein generative adversarial network with gradient penalty adversarial learning for providing a one-hot class vector to the discriminator network. To synthesize low-frequent and unknown attacks, the SAVAER's decode is used.

Due to the complexity of a SDN/NFV-based Agriculture 4.0 environment, the intrusion network samples are overwhelmed by a large number of normal samples. To resolve this issue, the work by Jiang *et al*. [72] can be adopted, in which the authors introduced a network intrusion detection algorithm. The proposed work is combined with a deep hierarchical network where one-side selection is used to reduce the noise of samples in the majority category. The performance evaluation showed that the proposed work can achieve an accuracy of 83.58% and 77.16% on the NSL-KDD and UNSW-NB15 dataset, respectively. To combat malware spread in IoT-based networks, Guizani and Ghafoor [73] designed a software-based architecture for network function virtualization. For predicting malware attacks, the proposed architecture used a machine-learning recurrent neural network long short term memory (RNN-LSTM) model. The BoT-IoT Dataset is used on experimentations, in which the tests showed the number of hosts decreased as a result of the RNN-LSTM identification of the malware attack.

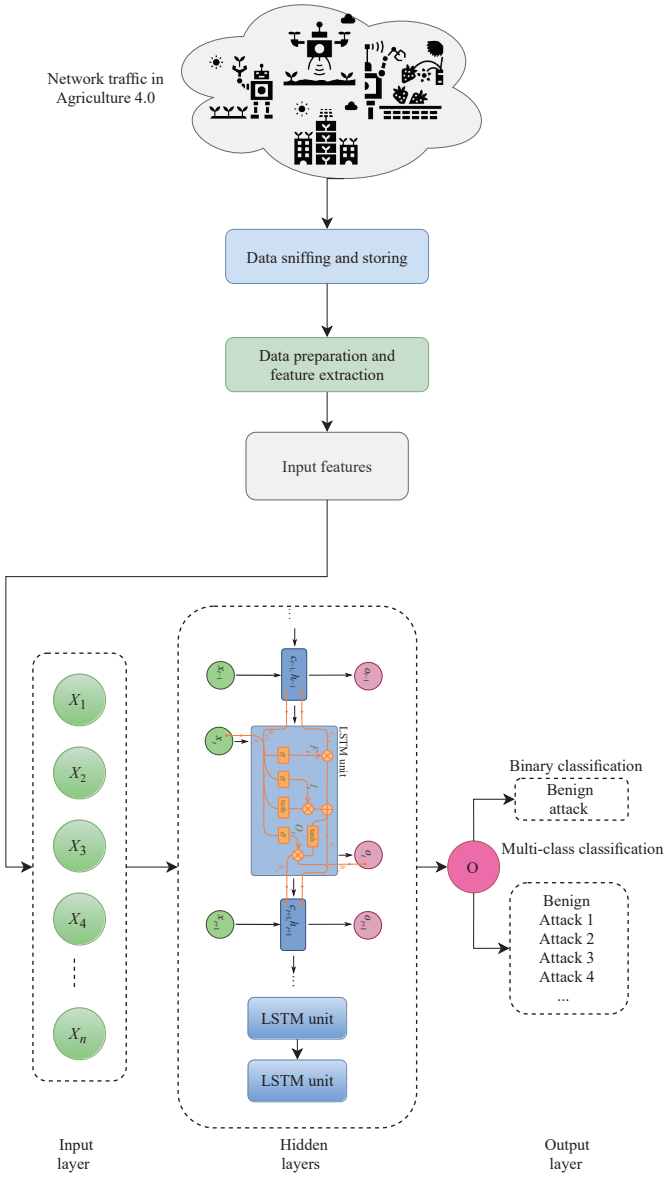

Fig. 8. Deep learning-based IDS for Agriculture 4.0.

D. IDS for Drones-Enabled Agriculture 4.0

Due to the integration of 5G systems in the emerging smart city idea, the internet of drones (IoD) has emerged as a new research field of "drone-to-drone communication (D2D)" for the Agriculture 4.0 [74]. The use of several UAVs (i.e., UAV swarm) collaborating to reach a specific goal in agriculture has increased productivity and reduced operational efforts [75], [76]. Nevertheless, these systems are vulnerable to cyber threats where an attacker can exploit them by causing significant damage such as taking control over them, disrupting operations, or stealing carried shipments [77]. Therefore, ensuring system security is becoming more and more crucial, especially in dynamic and decentralized drone-to-drone networks. Hence, finding an IDS for Agriculture 4.0 is still highly desirable. The IDSs for Drones-enabled Agriculture 4.0 can be categorized into two classes, namely, 1) Deep learning techniques and 2) Hierarchical approaches.

1) Deep Learning Techniques: Huang and Lei [78] designed

an imbalanced generative adversarial network-based IDS, named IGAN-IDS, for ad-hoc networks, which can be applied for drone-based Agriculture 4.0. Specifically, the IGAN-IDS is composed of three different units, namely, a deep neural network, an imbalanced generative adversarial network, and feature extraction component. To convert raw network properties into feature vector values, the IGAN-IDS adopts a feed-forward neural network, while the imbalanced generative adversarial network is used to generate new samples expressed in the latent space. The deep neural network is used for classifying attacks against drones-based Agriculture 4.0 using convolutional and fully-connected layers. Three datasets, namely, the CICIDS2017 dataset, UNSW-NB15 dataset, and NSL-KDD dataset, are used on the performance evaluation. The results showed that the multilayer perceptron achieved an accuracy of 78.32% on UNSW-NB15 and 78.97% on NSL-KDD.

To acheive the best performance in drone-based Agriculture 4.0, Wang *et al*. [79] can be applied to select optimal features. Specifically, the work used multilayer perceptions combined with sequential feature selection. The results of performance evaluation on the ISOT dataset and ISCX dataset showed that the proposed work can get the best test accuracy of 99.62%. However, based on encrypted network traffic analysis, Sciancalepore *et al*. [80] introduced a robust solution, named PiNcH, which can be used to detect the presence of a drone in Agriculture 4.0. The PiNcH solution is efficient in the presence of both heavy packet loss and evasion attacks. To boost the performance of IDS, Alzubi *et al*. [81] proposed a modified feature selection algorithm, named MBGWO, which is based on binary grey wolf optimization. The results of the performance evaluation on the NSL-KDD dataset showed that the MBGWO algorithm can achieve an accuracy of 99.22%, a detection rate of 99.10%, a false positive rate of 0.006% using only 14 features.

To select both feature subset and hyperparameters in one process, Elmasry *et al*. [82] introduced a double particle swarm optimization-based algorithm, which can be applied for drone-based Agriculture 4.0. In order to investigate performance differences, the authors used three deep learning models, namely, Deep Belief Networks, Long Short-Term Memory Recurrent Neural Networks, and deep neural networks (DBN). Two common IDS datasets, namely, NSL-KDD and CICIDS2017, are used in evaluation performance where the results showed that the proposed algorithm can reduce false alarm rate 1% to 5% and increase detection rate by 4% to 6%.

2) Hierarchical Approaches: Al Qurashi *et al*. [83] proposed an IDS architectural approach for resilient intrusion detection in ad-hoc networks, which can be applied for drone-based Agriculture 4.0. Abhishek *et al*. [84] designed a hybrid IDS in clustered IoT networks, which is deployed at the trusted node. The proposed system can be applied for drone-based Agriculture 4.0, where an Agriculture 4.0 network consists of one access point and a set of drone devices. To categorize the relay as either malicious or non-malicious, the proposed system performs a binary hypothesis test with the two hypotheses, namely, where i) the Relay is affected and compromising the packets and ii) the Relay is not affected and is in normal operation. Besides, based on the number of unicast packets dropped by the IoT devices, the proposed system can identify any adversary that affects the downlink unicast packet. By exchanging and sharing data, collaborative IDSs can improve the performance of a single detector in drone-based Agriculture 4.0. Based on the deep learning approach, Villamizar *et al*. [85] designed an accurate people detection approach, named WatchNet++, for detecting attacks in video surveillance, which can be applied for drone-based Agriculture 4.0. Li *et al*. [86] investigated the use of disagreement-based semi-supervised learning in collaborative IDSs. The performance evaluation on the DARPA dataset and a real dataset showed that the proposed system could outperform traditional supervised learning in terms of detection rate and as well as false alarm reduction.

E. IDS for Autonomous Tractors-Enabled Agriculture 4.0

In the last few decades, the implementation of autonomous tractors in Agriculture 4.0 has experienced a rapid growth [87].

Given the fact that autonomous vehicles operate on large interconnected networks, there is an increased risk of security and privacy measures [88]. However, the use of IDSs can minimize or mitigate such risks. The IDSs for Autonomous tractors-enabled Agriculture 4.0 can be categorized into four classes, namely, 1) Deep learning techniques, 2) Data analytics and statistical techniques, 3) Flow-based intrusion detection, and 4) Distributed collaborative intrusion detection. IDS for Autonomous tractors in Agriculture 4.0 is presented in Fig. 9.

1) Deep Learning Techniques: Based on a deep convolutional neural network, Song *et al*. [89] developed a system for intrusion detection for in-vehicle networks, which can be applied for autonomous tractors-based Agriculture 4.0 against cyber-attacks, such as denial-of-service and spoofing attacks. van Wyk *et al*. [90] combined a convolutional neural network with a well-established method of detecting anomalies, and Kalman filtering with an X2-detector for anomaly detection and identification in automated vehicles. van Wyk *et al*. [90] combined a convolutional neural network, named CNN-KF, with a well-established anomaly detection method, and Kalman filtering with an X2-detector for anomaly detection and identification in automated vehicles. The input to the convolutional neural network is a series of "images" from the connected and automated vehicles, and then classify these images as anomalous or normal. The numerical experiments demonstrated that the CNN-KF can detect anomalies and identify their sources with a high F1 score, sensitivity, and accuracy.

2) Data Analytics and Statistical Techniques: To extract more optimized and strongly correlated features, Ieracitano *et al*. [91] proposed a statistical analysis and autoencoder driven intelligent IDS, which is based on the combination of data analytics and statistical techniques. The proposed system can be applied for autonomous tractors-based Agriculture 4.0. The performance evaluation using the benchmark NSL-KDD dataset (i.e., binary-classification (Normal, Abnormal) and

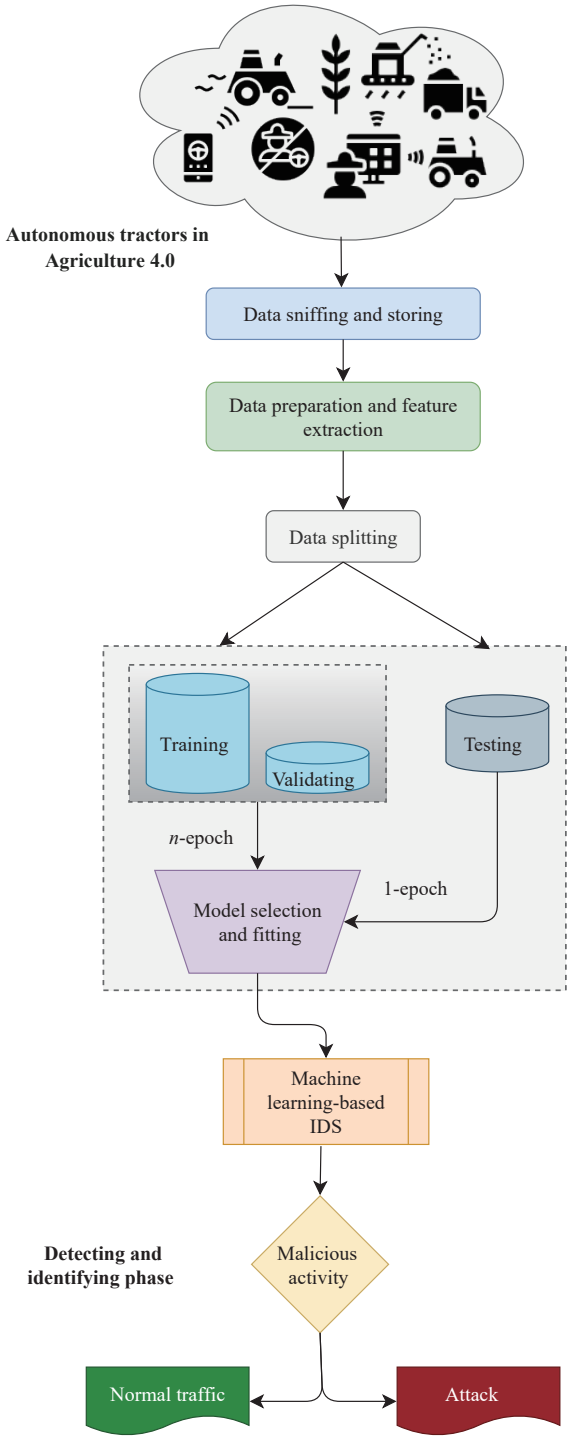

Fig. 9. IDS for autonomous tractors-enabled Agriculture 4.0.

multi-classification (Normal, Dos, R2L, Probe)) showed that the proposed system is efficient in terms of precision, recall, F1 score, and accuracy. Through payload analysis of network traffic, Vidal *et al*. [92] proposed an enhanced payload analyzer, named EsPADA, which takes advantage of the Bloom filtering data structure paradigm and the N-gram methodology. During the training stage, according to features extracted by N-gram, both normal and adversarial models are constructed. Both DARPA'99 and UCM 2011 datasets are used in the performance evaluation, where the results showed that the EsPADA can resist the disguised attacks with a lower accuracy reduction (4.86%).

To optimize processes of cyber security in large distributed systems, Vieira *et al.* [93] designed an autonomic intrusion detection and response system, which can be applied for autonomous tractor-based Agriculture 4.0. To evaluate the proposed system, the authors used a proof-of-concept implementation in two scenarios: i) virtual machines running on a private cloud, and ii) virtual machines running on Amazon public cloud. The results showed that the proposed system is effective within acceptable timeframes for both private and public cloud experiments. Therefore, to detect data integrity attacks autonomous tractors-based Agriculture 4.0, the work by Benisha and Ratna [94] can be applied, where the authors designed an IDS, named DI-EIDS, which is based on three techniques, including, the Deviation forest, Grey Wolf Optimization, and Black forest classifier. To obtain the best training data, the DI-EIDS adopts a Black forest classifier, while Grey Wolf Optimization is implemented for sampling ratio optimization. The UNSW-NB15 dataset is used on evaluation performance, where the results showed that the DI-EIDS can achieve higher performance in accuracy and FAR.

3) Flow-Based Intrusion Detection: To improve the detection rate of minority classes, Zhang *et al*. [95] designed a flow-based intrusion detection model, named SGM-CNN, which uses a combination of synthetic minority over-sampling technique (SMOTE) and under-sampling for clustering based on Gaussian Mixture Model. Both UNSW-NB15 and CICIDS2017 datasets are used in the performance evaluation, where the experimental results showed that the SGM-CNN model can achieve detection rates of 99.74% and 96.54% on the UNSW-NB15 dataset using binary classification and multiclass classification, respectively. In addition, the SGM-CNN model can achieve a detection rate of 99.85% on the CICIDS2017 dataset for 15-class classification. Therefore, Liu *et al*. [96] proposed a web IDS based on the combination of feature analysis and support vector machine optimization to find a kernel function, which can be applied for autonomous tractor-based Agriculture 4.0. The HTTP DATASET CSIC 2010 dataset is used in the experiments, in which the results showed that the proposed system has better detection capability.

4) Distributed Collaborative Intrusion Detection: The characteristics of high mobility and rapid topology change of autonomous tractors in Agriculture 4.0 makes it vulnerable to various malicious attacks. To enable data collection, analysis, and tracking between vehicles, Zhou *et al*. [97] proposed a distributed collaborative IDS, named DCDIV. To ensure a reliable and stable communication link, the DCDIV system adopts a reputation-based cooperative communication method. The Venis tool is used as an experimental scenario for simulating the traffic and network environment of vehicle communications. The results showed that the DCDIV system can achieve a faster attack detection rate, lower false alarm rate, and higher detection rate. Therefore, based on random forest feature selection, Li *et al*. [98] designed an auto-encoder IDS, named AE-IDS, which can be applied for autonomous tractors-based Agriculture 4.0. The AE-IDS

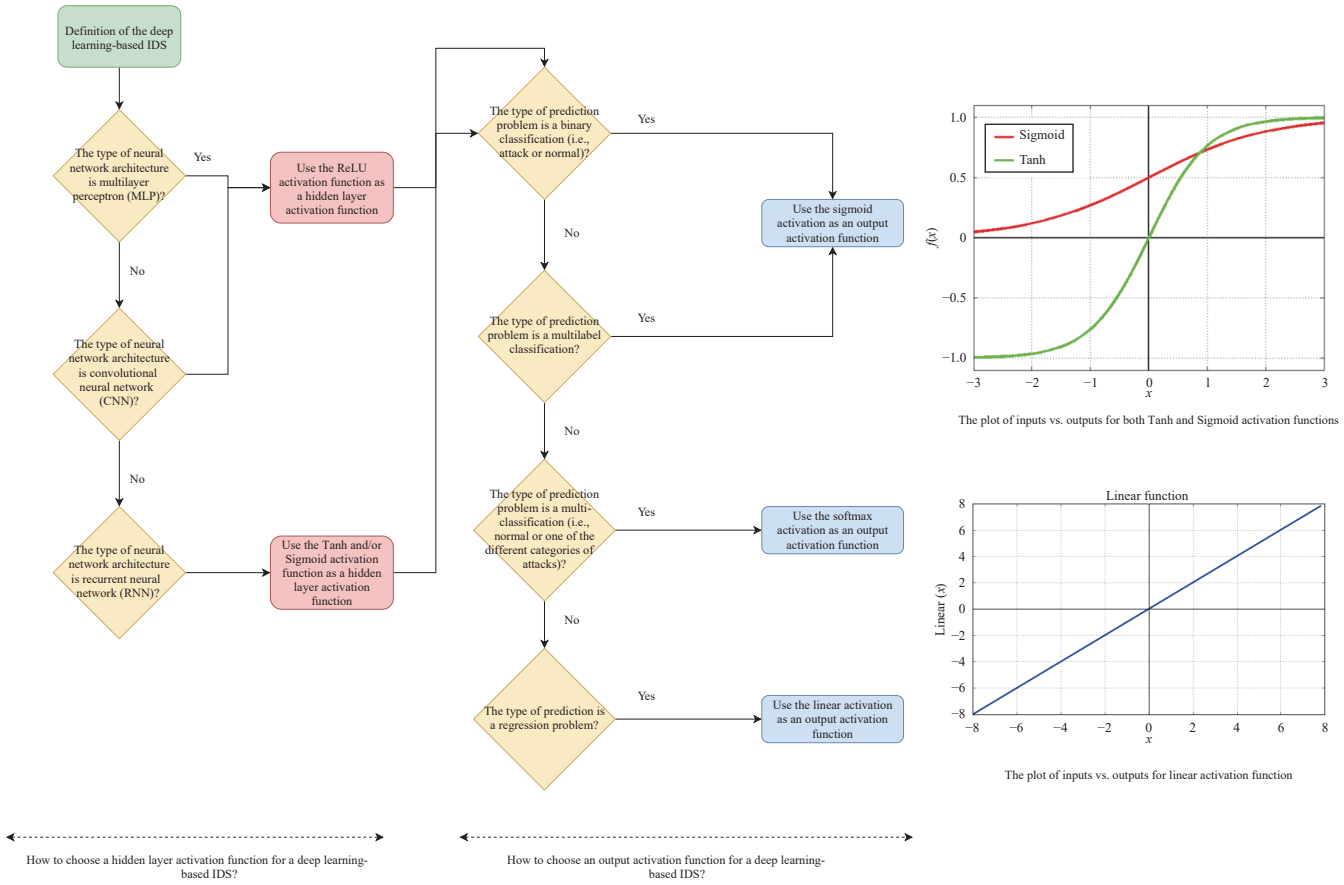

Fig. 10. Flow diagram for deciding how to choose the type of hidden layer activation function and output activation function for a deep learning-based IDS.

architecture is made up of four main connected modules, namely: i) Data Preprocessing, which consists of dividing the dataset into a training set and a test set by proportion, ii) Feature Selection, which consists of selecting the most significant feature using the random forest algorithm, iii) Feature Grouping, which consists of applying the affinity propagation clustering algorithm to group the average features and get various feature subsets, and iv) Anomaly Detection, which consists of the combination of auto-encoder and K-means. The CSE-CIC-IDS2018 dataset is used as an experimental environment, in which the results showed that the AE-IDS can perform better than some popular batch/offline methods.

F. IDS for IoT Devices-Enabled Agriculture 4.0

The IDSs for IoT devices-enabled Agriculture 4.0 can be categorized into four classes, namely, 1) Hierarchical approaches, 2) Deep learning techniques, 3) Hybrid machine learning, and 4) Artificial bee colony (ABC) model-based.

1) Hierarchical Approaches: To secure IoT devices and detecting cyberattacks in the IoT environment, Ferrag *et al*. [99] developed a system for intrusion detection, named RDTIDS, which is based on rules and decision tree algorithms. The RDTIDS uses two steps: training and testing. The training step consists of training the three classifiers under a hierarchical model, while the testing step consists of the classification of data as either benign or as a specific type of attack. Therefore, the RDTIDS is located in the fog computing layer in a three-tier fog computing architecture. Both the CICIDS 2017 dataset and the Bot-IoT dataset are used for the experiments, in which the results showed that the RDTIDS can achieve the highest true negative rate with 98.855%.

2) Deep Learning Techniques: The activation functions in the hidden layer as well as in the output layer are an essential feature of the design of a Deep learning-based IDS. Specifically, the selection of the activation function in the hidden layer determines the ability of the network model to learn the training data set, while in the output layer, it determines the specific type of predictions that the model can perform [100]. Fig. 10 presents a procedure that could be applied as an initial check for deciding how to choose the type of hidden layer activation function as well as an output activation function for a deep learning-based IDS. Almiani *et al*. [101] introduced an artificially full-automated IDS for IoT devices, which can be applied for IoT devices-based agriculture. To catch specific types of attacks for IoT environments, the proposed system uses an architecture of cascaded filtering stages and adopts the deep multi-layered recursive neural networks. The proposed training algorithm uses the following stages: Feedforward computation, Backpropagation to the output layer, Backpropagation to the hidden layer, and Weights update. The NSL-KDD dataset is used as a training and testing benchmark dataset, in which the results showed detection rates of 98.27%, 97.35%, 64.93%, and 77.25%, for DoS attacks, Probe attacks, R2L attacks, and

U2R attacks, respectively. Li *et al*. [102] focused on the use of a convolutional neural network to overcome the problems of image security detection and adapt it to the present diversity of media types, and to build a high-speed, high-precision imaging type recognition system. The proposed scheme results in more than 93.75% classification accuracy across the general image library.

3) Hybrid Machine Learning: To perform feature clustering for efficient detection of intrusions, Aljawarneh and Vangipuram [103] introduced a Gaussian dissimilarity measure, named GARUDA, for anomaly detection in IoT networks. The performance evaluation is conducted on both KDD and NSL-KDD datasets, in which the results showed detection accuracies of 85.82%, 19.23%, 94.83%, 99.07%, and 98.38% for DoS, Probe, R2L, U2R, and Normal classes, respectively. Based on long short term memory recurrent neural networks, Jiang *et al*. [104] proposed a multi-channel intelligent attack detection system. The proposed system uses the following steps: Data preprocessing step, Multi-channel processing step, and Voting step. The Data preprocessing step consists of providing high-quality data and extracting different types of features from the processed data. The Multi-channel processing step consists of generating classifiers by training neural networks. The Voting step consists of deciding whether the input data is an attack or not. The NSL-KDD dataset is used to train and test the multi-channel intelligent attack detection system, in which the results showed a higher accuracy of 98.94%.

Botnets are the most serious threats to cyber Agriculture 4.0, where, a large number of hosts are remotely controlled by attackers with zombie programs. To detect botnets, Wang *et al*. [105] proposed an automated botnet detection system, named BotMark, which can be applied in complex environments such as Agriculture 4.0. The BotMark system does not require any previous knowledge of botnets and utilizes a hybrid analysis of flow-based and graph-based traffic patterns. To characterize the behaviors of botnets, the BotMark system extracts from network traffic 15 flow-based features and 3 types of graph-based features. The experimental results showed that the BotMark system can reach a detection accuracy of 99.94%. Shafiq *et al*. [106] presented a feature selection metric approach called CorrAUC, together with a feature selection algorithm called Corrauc, that can be used in IDSs for securing IoT networks. The authors assessed the proposed approach by utilizing the Bot-IoT dataset and four separate ML algorithms. The results showed that the suggested method can achieve results greater than 96% on average. Al Shorman *et al*. [107] proposed an unsupervised evolutionary IoT botnet detection method, named GWO-OCSVM, based on the grey wolf optimization algorithm (GWO) and one-class support vector machine (OCSVM). The GWO algorithm is used to optimize the hyperparameters of the OCSVM. The experimental results showed that the GWO-OCSVM system with the NN-BaIoT dataset showed good performance in terms of the true positive rate and false-positive rate. To efficiently detect network intrusions, Hassan *et al*. [108] proposed a hybrid deep learning model using a long short-term memory network and a convolutional neural network. The UNSW-NB15 dataset is used as a publicly available big dataset for the performance evaluation of the hybrid deep learning model, which shows good performance, achieving 97.1% accuracy compared to traditional approaches.

4) Artificial Bee Colony (ABC) Model-Based: The Sybil attack can create serious threats to cyber Agriculture 4.0, where an adversary asserts various illegal identities by constructing or damaging the IoT nodes. To detect the Sybil attack in the IoT environment, Murali and Jamalipour [109] proposed a lightweight IDS based on the ABC model. The ABL model is used as an optimization technique for simulates the foraging behavior of honey bees. The simulation results showed that the average accuracy rate of the proposed IDS is 96.8%, 95.2%, and 94.8%, for type 1, type 2, type 3 attack, respectively. The type 1 attack consists of malicious nodes that will target one fixed region. The type 2 attack consists of malicious nodes that are scattered among the legitimate nodes, while the type 3 attack consists of Sybil nodes under mobility and is distributed among the network. To address intrusion detection in supervised problems, the work by Lopez-Martin *et al*. [110] can be applied, in which the authors use reinforcement learning to network intrusion detection using two datasets, namely NSL-KDD and AWID datasets. The study evaluated the IDS model regarding the use of the following four deep reinforcement learning methods: actor-critic (AC), policy gradient (PG), double deep Q-network (DDQN), and deep Q-network (DQN). The DDQN algorithm showed good results compared to other deep reinforcement learning algorithms.

G. IDS for Industrial Agriculture 4.0

Critical infrastructure is essential for the good governance of society. This is why the exposure of such infrastructures to extreme events has a significant impact on the resilience of society [111]. The industrial Agriculture 4.0 is complex and diverse which shows a low real-time performance for impersonation attacks. The IDSs for Industrial Agriculture 4.0 can be categorized into two classes, namely, 1) Hierarchical approaches and 2) Hybrid machine learning.

1) Hybrid Machine Learning: Liang *et al.* [112] proposed an industrial network IDS that can be applied for Agriculture 4.0. The proposed system uses a multi-feature data clustering optimization model to diagnose, restore, and rebuild. The performance evaluation using both NSL-KDD and KDDCU'99 datasets showed that the proposed system can reach an accuracy of 97.8%. Therefore, for secure smart factory-based ambient intelligence, Park *et al*. [113] introduced a machine learning and context-aware IDS. The proposed system consists of three phases, namely, i) Data capture and parsing phase, ii) Model building and inference phase, and iii) Threat visualization phase. For reconstruction and compensation of cyber attacks launched on industrial IoT systems, Farivar *et al*. [114] designed a hybrid intelligent-classic control approach. As an intelligent estimator for attack estimation, the proposed approach uses a neural network. To ensure system stability during attacks, the proposed approach adopts the nonlinear control theory.

2) Hierarchical Approaches: Industrial cyber-physical systems (ICPS), combining advanced communication, computing, and industrial process monitoring, are considered a core technology for Industry 4.0. Liu *et al*. [115] introduced a hierarchically distributed intrusion detection scheme to achieve all-round security protection of ICPSs. An adversary can launch a stealthy attack to manipulate the sensor readings secretly based on the knowledge of the physical model used by traditional IDS. To detect stealthy attacks on an industrial control system, Hu *et al*. [116] designed an intrusion detection approach using permutation entropy. Therefore, to study the proposal's efficiency in detecting stealthy attacks, a Matlab-Simulink environment is used, in which the results showed a good detection ability of the proposed permutation entropy-based intrusion detection against stealthy attacks on industrial control systems. To estimate the attack signal waveforms in an industrial control system, Miao *et al*. [117] proposed the following two types of estimators: nonlinear attack signal estimator (NASE) and linear attack signal estimator (LASE). However, the IDSs proposed in the industrial control systems can be classified into system model-based systems and traditional information technology-based systems.

H. IDS for Smart Grid-Enabled Agriculture 4.0

The smart grid-based Agriculture 4.0 is composed of a set of controllers, automation, and standard communication protocols, where they are interconnected over the Internet to control the production and distribution of energy to IoT devices in smart agriculture [21]. The IDSs for Smart Grid-enabled Agriculture 4.0 can be categorized into two classes, namely, 1) Reinforcement learning and 2) Collaborative intrusion detection.

1) Reinforcement Learning: Kurt *et al*. [118] designed a reinforcement learning approach for online cyber-attack detection, which can be applied for smart grid-based Agriculture 4.0. The proposed approach uses a direct mapping from observations to actions using two phases, namely, the training phase and the test phase. The training phase consists of training the defender with low magnitude attacks, while the test phase consists of detecting slight deviations of meter measurements. The IEEE-14 bus power system is used as a simulation environment, in which the results showed a high potential of reinforcement learning approaches in solving complex cyber security problems such as smart grid-based Agriculture 4.0.

2) Collaborative Intrusion Detection: To provide the best possible protection of Smart Grid ecosystems, Patel *et al.* [119] proposed a collaborative IDS, named IDPS, which can identify an attack based on three structural forms, namely, Centralized, Hierarchical, and Fully distributed. The IDPS uses three advanced components: a fuzzy logic risk manager, knowledge manager, and an autonomic manager. To classify the binary-class, triple-class, and multi-class cyber-attacks in the smart grid, Haghnegahdar and Wang [120] proposed an instruction detection system using a whale optimization algorithm and an artificial neural network. The proposed system uses the artificial neural network to achieve the minimum mean square error, while the Whale optimization algorithm is applied to initialize and adjust the weight vector.

Tables III and IV present a summary of IDSs for Agriculture 4.0. The classification of IDS solutions for Agriculture 4.0 is presented in Fig. 11.

IV. IDS Building Process and Public Datasets

The food industry has experienced a shift from disconnected, stand-alone, independent operations to heavily interconnected, dependent, and integrated operations, to improve the sector's efficiency [87]. As a result, network organizations find themselves in a highly efficient production system, with growing complexity, and increased exposure to potential risks. Connectivity in the agri-food chain involves the control of information assets, the transport of physical goods and services, and other intangible assets, as shown in Fig. 12. Increasingly, this control has become ubiquitous and pervasive throughout Agriculture 4.0, making it even more difficult to secure all of the sector's resources. In this section, we provide the IDS building process for Agriculture 4.0.

A. Agriculture 4.0 Data Sources

Embedding emerging technologies provides smarter management of agri-food supply chains, as it can combine diverse patterns of independent data analysis, historical data repositories, and real-time data traffic [127]. Both real-time data and automated data processing tools offer new ways to respond more quickly to changing conditions in Agriculture 4.0. The activities associated with each agricultural component are automatically integrated into the food chain through emerging technologies, from farm to fork, as shown in Fig. 12. Some of these components have separate data sources that need to be used. These data not only must be present, but they must also work in balance across all systems. Some of the core components in Agriculture 4.0 are:

1) Smart Farming Systems: These systems are designed by integrating advanced technologies into existing farming operations, such as intelligent crop/livestock monitoring, smart water management, disease management, smart harvesting, etc., to improve the quality and efficiency of agricultural production. It includes different types of sensors, actuators, unmanned aerial/ground vehicles, smart agricultural machinery, and so on, focusing on linking objects in the IoT-based smart farm. While monitoring, performing agricultural tasks, and processing farm-related data, via deployed intelligent devices.

2) Transportation Services: These services are in charge of the flow of agricultural products from one location to another, starting at the beginning of the supply chain and ending at the customer's kitchen. It includes different types of smart sensors, GPS kits, Internet of vehicles (IoV) communications, where vehicles communicate among themselves and with public networks through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) interfaces. They permit the real-time collection and sharing of sensitive information on road conditions and agricultural payloads status.

3) Storage Entities: These entities are in charge with storage operations. Cold storage systems are equipped with monitoring systems that can track changes over time in the

TABLE III
SUMMARY OF IDSS FOR AGRICULTURE 4.0

System	Year	Network model	IDS model	Technique	Validation	Attacks	Pros (+)	Cons (–)
Zhong *et al*. [121]	2021	IoT Servers	machine learning system	- Detect/identify anomalous data and flow in IoT environement	Network intrusion detection dataset	Attacks on two datasets, namely KDD99 and ADFA-LD datasets	+ Good performance for malicious behavior detection and recognition	- RoC curve is not reported
Gill *et al*. [35]	2020	A cloud environment divided into four layers: Attack Layer, Security Layer, Back-end Layer, and End-user Layer.	Game theoretic model	- Signature-based- Anomaly-based- Honeypot	Simulation environment in MATLAB	Attacking the cloud servers with regular or sophisticated attacks	+ Reduce energy consumption by the defender system	- The threat model is limited
Rabbani *et al*. [36]	2020	A cloud-based environment divided into different levels of sub-behaviors	A hybrid machine learning system	- A particle swarm optimization-based probabilistic neural network	Network intrusion detection dataset	Malicious behaviors such as Worms, Backdoors, Fuzzers, etc.	+ Good performance for malicious behavior detection and recognition	- Requires a comparison with deep learning techniques
Kushwah and Ranga [37]	2020	A cloud infrastructure with a detector attached	A machine learning system	- Voting extreme learning machine	Network intrusion detection dataset	DDoS attacks	+ High accuracy of 99.18% with the NSL-KDD dataset and 92.11% with the ISCX dataset	- There are only three metrics used (i.e., accuracy, sensitivity, and specificity)
Aldribi *et al*. [38]	2020	A cloud environment	A hypervisor-based IDS	- Online multivariate statistical change analysis	Network intrusion detection dataset	Input validation, authentication breach, backdoors, DoS, etc.	+ The overall detection rate = 96.23% and False positive rate = 7.56%	- IoT data is not considered
Almogren [41]	2020	Edge-of-Things computing	A hybrid machine learning system	- Malicious activity detection model	Network intrusion detection dataset	Nine attacks: worms, DoS, misuses, etc.	+ The deep belief network has the best overall performance as compared to the artificial neural network and support vector machine	- The network model is not defined
Almiani *et al*. [101]	2020	IoT devices connected with fog node	An artificially full-automated IDS	- Artificial intelligence-based approach	Network intrusion detection dataset	DoS attacks, Probe attacks, R2L attacks, and U2R attacks	+ The results show detection rates of 98.27%, 97.35%, 64.93%, and 77.25%, for DoS attacks, Probe attacks, R2L attacks, and U2R attacks, respectively	- The IoT data traffic is not considered
Huang and Lei [78]	2020	Dynamic and decentralized ad-hoc networks.	An imbalanced generative adversarial network towards IDS	- Artificial intelligence-based approach	Network intrusion detection dataset	Attacks on three datasets, namely, CICIDS2017 dataset, UNSW-NB15 dataset, and NSL-KDD dataset	+ Achieves at least 1% and 6% improvement on Precision and Recall, respectively	- The IoT traffic data are not considered
Abdulqadder *et al*. [60]	2020	SDN/NFV enabled cloud of 5G networks	A multilayered intrusion detection and prevention system	- Artificial intelligence-based approach	Simulation environment in NS3	Host location hijacking, control plane saturation, DDoS attack	+ Efficient in terms of security between switches and controllers	- IoT data is not considered and the RoC curve is not reported
Zhou *et al*. [97]	2020	Connected and automated vehicles	Distributed collaborative IDS	- Identify betray attacks in VANET	Simulation environment in Venis tool	Spoofing attacks, Black-Hole attacks, Gray-Hole attacks, and Denial-of-Service attacks	+ Can achieve a faster attack detection rate, lower false alarm rate, and higher detection rate	- The IoT data traffic is not considered
Ngo *et al*. [57]	2020	A SDN network architecture which includes infrastructure layers, control, and application	A heterogeneous hardware-based network IDS	- Signature-based- Anomaly-based	Parallel hardware platforms	SYN flood packets	+ Intrusion detection engines are 14× faster under GPU compared to CPU	- The threat model is limited
Tian *et al*. [40]	2020	Distributed edge devices	A distributed deep learning system	- Analyzing URLs	Network intrusion detection dataset	SQL injection, XSS, and command injection	+ The accuracy fluctuates by approximately 0.955	- RoC curve is not reported

TABLE IV
SUMMARY OF IDSS FOR AGRICULTURE 4.0 (CONTINUED)

System	Year	Network model	IDS model	Technique	Validation	Attacks	Pros (+)	Cons (–)
van Wyk *et al*. [90]	2020	Connected and automated vehicles	A convolutional neural network-based IDS model	- Detect/identify anomalous sensor values	Simulation to generate datasets	Attacks against connected and automated vehicles	+ Can detect anomalies and identify their sources with high F1 score, sensitivity, and accuracy	- The IoT data traffic is not considered
Murali and Jamalipour [109]	2020	The routing protocol for low power and lossy networks (RPL) in IoT networks	A lightweight IDS	- Detect/identify Sybil attack	Simulation environment in Cooja under contiki OS	Three types of Sybil attack	+ Average accuracy rate of the proposed IDS is 96.8%, 95.2%, and 94.8%, for type 1, type 2, type 3 attack, respectively	- RoC curve is not reported
Jiang *et al*. [104]	2020	IoT data	A multi-channel intelligent attack detection system	- Detect/identify anomalous data	Network intrusion detection dataset	DoS attacks, Probe attacks, R2L attacks, and U2R attacks	+ A higher accuracy of 98.94% compared to Bayesian or SVM classifiers	- DDoS attacks are not considered
Kurt *et al*. [118]	2019	Smart Grid	A reinforcement learning-based attack detection scheme	- Detect/identify anomalous data	Simulation environment in IEEE-14 bus power system	False data injection	+ A high potential of reinforcement learning approaches in solving complex cyber security problems	- The IoT data is not considred
Patel *et al*. [119]	2017	Smart Grid ecosystems	A collaborative IDS	- Detect/identify anomalous data	Simulation environment in Network Simulator NS-2.33	Force, command injection and brute	+ Good detection rate with low false positive alarms	- The botnet attack is not considred
Potluri and Diedrich [122]	2016	A cloud environment	An accelerated deep learning model	- Artificial intelligence-based approach	Network intrusion detection dataset	Attacks on NSL-KDD dataset	+ Efficient in terms of fast detection	- Not relevant for the detection of new attacks.
Lin *et al*. [123]	2015	A cloud environment	A machine learning system	- Artificial intelligence-based approach	Network intrusion detection dataset	Attacks on KDD CUP 99 dataset	+ High computational efficiency for the time of classifier training and testing	- The threat model is limited
Gao *et al*. [124]	2014	A cloud environment	A deep hierarchical learning model	- Deep Belief Networks model	Network intrusion detection dataset	Attacks on KDD CUP 99 dataset	+ Demonstrates that the performance of Deep Belief Networks model is better than that of SVM and ANN.	- Not relevant for the detection of new attacks.
Tesfahun and Bhaskari [125]	2013	A cloud environment	A machine learning system	- Detect/identify anomalous data	Network intrusion detection dataset	DoS attack, Probing attack, R2L attack and U2R attack	+ Good detection rate with low false positive alarms	- Not relevant for the detection of new attacks.
Li *et al*. [126]	2012	A cloud environment	A hybrid machine learning system	- Detect/identify anomalous data	Network intrusion detection dataset	DOS, R2L, U2R, probing, surveillance and other probing.	+ The accuracy was 98.6249% in 10-fold cross validation and the average Matthews correlation coefficient (MCC) achieves 0.861161.	- The threat model is limited

status of the stored agricultural products, warning and alerting management as soon as something seems to be wrong. It comprises several types of smart sensors, such as temperature and humidity sensors.

4) Food Processors: These systems both prepare fresh food for the market as well as manufacture prepared agricultural products. It is composed of a relatively large and diverse group of companies that make a product. They also use agricultural raw materials or sub-assemblies manufactured by different producers to develop their products. Using IoT-enabled equipment, it is possible to manage a wide range of quality control operations. For example, manufacturers can monitor production volumes and temperatures for different commodities, pressure conditions, and labeling products.

5) Distributors: These services would usually consist of an entity that acquires large inventories of products that it purchases from producers and sells to consumers. Distributors fulfill the "Time and Place" requirement for the customer by delivering the products whenever and wherever the customer desires.

6) Retailers: These entities hold stocks with smaller quantities for sale to the public. They also keep track of the customers' preferences and demands.

B. Public Datasets

In this sub-section, we examine 12 selected datasets made available since 1999. Table V present a summary of public datasets and selected IDSs, that can be used in Agriculture 4.0.

1) KDD99 Dataset: This dataset was the most widely used dataset for the evaluation of IDSs [48]. This dataset is created by Stolfo *et al*. [142] using data collected in the DARPA'98 IDS evaluation program [143], which is approximately 4

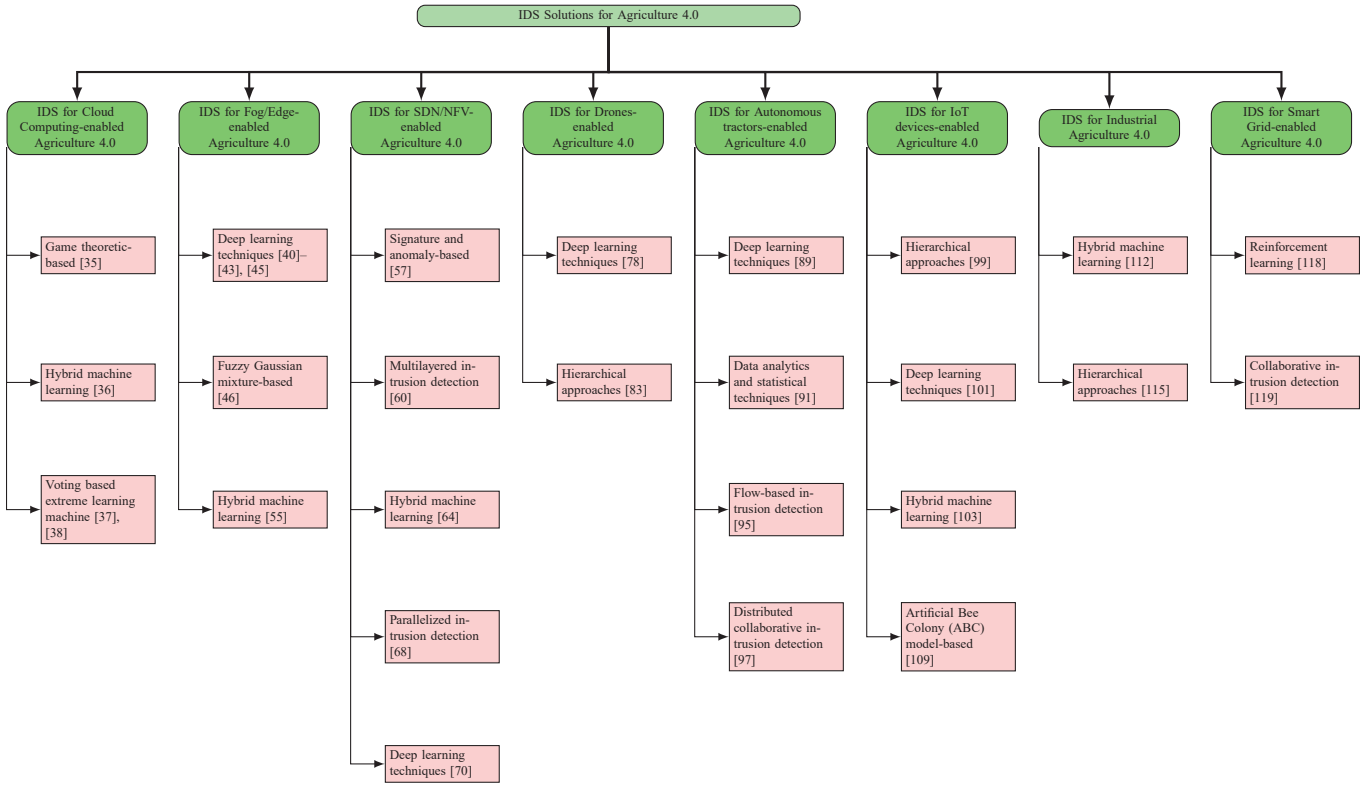

Fig. 11. Classification of IDS Solutions for Agriculture 4.0.

gigabytes of raw Tcpdump data packed from 7 weeks of network traffic, that can be managed in about 5 million connection records [129]. KDD99 consists of approximately 4 900 000 vectors with 23 types of attacks divided into four major attack categories, namely, remote to local (R2L), probing (PRB), denial of service (DoS), and user to root (U2R).

2) UNSW-NB15 Dataset: This dataset created by the IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) [128], in order to provide a hybrid of real modern normal activities and synthetic attack patterns. The Tcpdump tool was selected to capture 100 GB of raw traffic. It contains approximately 2 540 044 data instances with 49 features and 9 types of attack in total, including, Worms, Reconnaissance, Shellcode, Generic, Exploits, and DoS.

3) NSL-KDD Dataset: This dataset was suggested by [129], to resolve certain problems inherent in the KDD'99 dataset, such as the lack of a precise definition of the attacks. It contains two datasets KDDTrain+ (125 973 records) and KDDTest+ (22 544 records), which are generated from the KDD'99 data set. There are four major attack categories, namely, R2L, PRB, DoS, and U2R.

4) ISCX Dataset: This dataset was created by the Information Security Centre of Excellence (ISCX) [130], consists of 7 days of network activity, and is based on four types of attacks, namely, Distributed Denial of Service, Brute Force SSH, HTTP Denial of Service, and Infiltrating the network.

5) ISOT Cloud Intrusion Dataset: This dataset consists of an aggregation of more than 8 terabytes collected synchronously from an OpenStack cloud production environment [131]. The dataset contains web vulnerabilities scanning, dictionary/brute force login,directory/path traversal, cross-site scripting, SQL injection, fuzzers, and HTTP flood DOS attack types.

6) HTTP Dataset CSIC: This dataset was created by the Information Security Institute of CSIC (Spanish Research National Council) [132]. It holds traffic that was generated for an e-commerce web application, and contains more than 36 000 normal requests and 25 000 anomalous requests with some types of attacks, such as parameter tampering, CRLF injection, XSS (Cross-site scripting), SQL injection, etc.

7) Kyoto 2006 Dataset: The raw traffic data was collected by honeypot systems deployed at Kyoto University [133]. It contains 14 features were extracted based on KDD Cup 99 data set and 10 additional features.

8) CICIDS 2017 Dataset: This dataset was created by the Canadian Institute for cyber security [134]. It has 80 features associated with every Netflow record written in CSV format, which facilitates its import into a machine learning package. It contains the most common attacks based on the 2016 McAfee report, such as Scan, Web based, Brute force, DDoS, Heartbleed, Infiltration, and Bot.

9) Industrial Control System Cyber Attack Dataset: This dataset contains five datasets, including, i) Power System Datasets, ii) Gas Pipeline Datasets, iii) Gas Pipeline and Water Storage Tank, iv) New Gas Pipeline, and v) Energy Management System Data [135].

10) AWID Dataset: For data collection, [136] created an actual laboratory that realistically reproduces a typical SOHO infrastructure. Several mobile and stationary STAs were used as valid clients of the network, while a unique mobile attacker

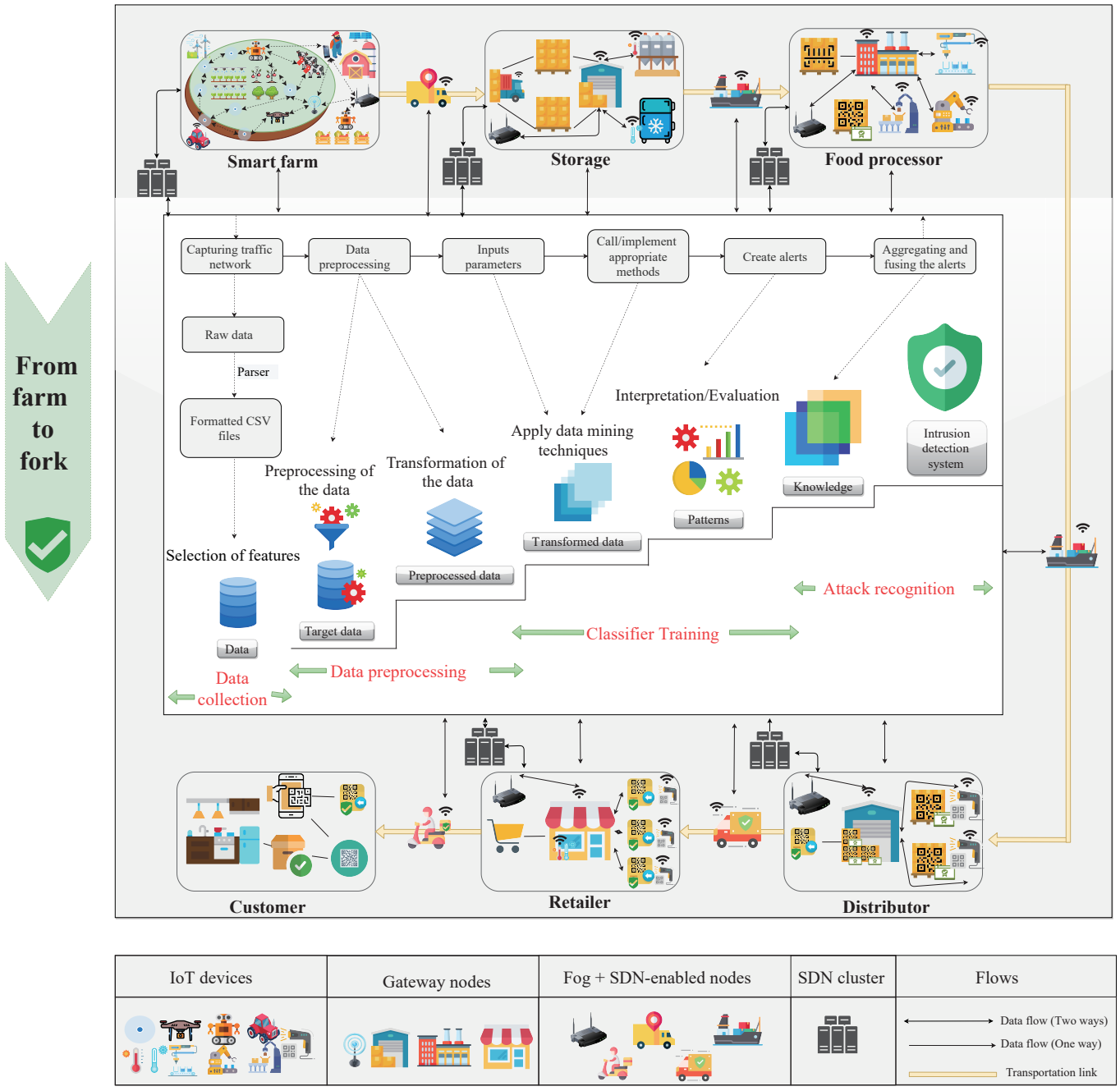

Fig. 12. IDS-secured smart supply chain in Agriculture 4.0.

triggered various attacks. It contains a vector of 155 attributes, and each attribute has numeric or nominal values. There are three types of attacks, including, Injection, Flooding, and Impersonation.

11) CSE-CIC-IDS2018 Dataset: This dataset was created by the Canadian Institute for cyber security [134]. It contains a collection of web application attacks, brute force attacks, and last updated attacks, such as DDoS+PortScan, Botnet attacks, Infiltration attacks, Web attacks, DoS attacks, and Bruteforce attacks.

12) Bot-IoT Dataset: This dataset was generated by creating a realistic network environment in the Cyber Range Lab at the UNSW Canberra Cyber [137]. It contains more than 72.000.000 records with the following attacks, OS and Service Scan, DDoS, DoS, Data exfiltration attacks, and Keylogging.

13) TON_IOT Dataset: For the purpose of collecting and analysing heterogeneous data sources from the IoT and industrial IoT (IIoT), the IoT Lab of the UNSW Canberra Cyber, the School of Engineering and Information Technology (SEIT), UNSW Canberra at the Australian Defence Force Academy (ADFA) created a dataset named TON_IOT [138]. The testbed was implemented using multiple virtual machines comprising multiple operating systems to accommodate the interconnectivity among the three layers of IIoT, Cloud, and Edge/Fog systems. A variety of AI-based cyber security applications can be validated and tested with this data set, such as IDSs, threat intelligence, malware detection, fraud detection, and others.

TABLE V
PUBLIC DATASETS FOR CYBER SECURITY INTRUSION DETECTION IN AGRICULTURE 4.0

Datasets	Year	IDS system	Network model	Focus
KDD99 dataset [48]	1999	Haider *et al*. [46]	Edge computing	Train and validate fuzzy Gaussian mixture-based correntropy models
		Lv *et al*. [66]	SDN/NFV	Validate the effectiveness of an extreme learning machine with a hybrid kernel function to build an intrusion detection approach
		Liang *et al*. [112]	Industrial network	Evaluations and analysis of an industrial network IDS
		Chellammal and Malarchelvi [68]	SDN/NFV	Evaluate the performance evaluation of a parallelized intrusion detection architecture
UNSW-NB15 dataset [128]	2015	Rabbani *et al*. [36]	Cloud computing	Evaluate the performance of a hybrid machine learning approach in a cloud computing environment
		Almogren [41]	Edge computing	Analyze a malicious activity detection model for edge-of-things computing
		Hassan *et al*. [108]	IoT devices	Validate the effectiveness of a hybrid deep learning model
		Zhang *et al*. [95]	Autonomous tractors	Train and validate a flow-based intrusion detection model
		Lv *et al*. [66]	SDN/NFV	Validate the effectiveness of an extreme learning machine with a hybrid kernel function to build an intrusion detection approach
NSL-KDD dataset [129]	2009	Kushwah and Ranga [37]	Cloud computing	Evaluate the performance of a voting extreme learning machine for detecting DDoS attacks in a cloud computing environment
		Hosseini and Zade [55]	Edge computing	Evaluate the performance of a hybrid intrusion detection system that uses an artificial neural network
		Wu *et al*. [43]	Edge computing	Analyze a network IDS based on semantic Re-encoding and deep learning
		Kur and Singh [54]	Edge computing	Evaluate the performance of a hybrid intrusion detection system of unknown web attacks
		Chellammal and Malarchelvi [68]	SDN/NFV	Evaluate the performance evaluation of a parallelized intrusion detection architecture
ISCX dataset [130]	2012	Kushwah and Ranga [37]	Cloud computing	Evaluate the performance of a voting extreme learning machine for detecting DDoS attacks in a cloud computing environment
		Wang *et al*. [79]	Internet of Drones	Validating the effectiveness of a dynamic multilayer perceptrons-based DDoS attack detection system
ISOT Cloud Intrusion dataset [131]	2020	Aldribi *et al*. [38]	Cloud computing	Evaluate the performance of a hypervisor-based IDS for detecting anomalous cloud behavior
HTTP Dataset CSIC [132]	2010	Tian *et al*. [40]	Edge computing	Evaluate the performance of web intrusion detection
		Liu *et al*. [96]	Autonomous tractors	Evaluate the performance of web intrusion detection
Kyoto 2006 dataset [133]	2006	Chellammal and Malarchelvi [68]	SDN/NFV	Evaluate the performance evaluation of a parallelized intrusion detection architecture
CICIDS 2017 dataset [134]	2018	Kur and Singh [54]	Edge computing	Evaluate the performance of a hybrid intrusion detection system of unknown web attacks
		Zhou *et al*. [65]	SDN/NFV	Evaluate an IDS based on feature selection and ensemble classifier techniques
		Huang and Lei [78]	Dynamic and decentralized ad-hoc networks	Analyze an imbalanced generative adversarial network towards IDS
		Zhang *et al*. [95]	Autonomous tractors	Train and validate a flow-based intrusion detection model
		Khammassi and Krichen [69]	SDN/NFV	Evaluate a wrapper approach for feature selection
Industrial Control System Cyber attack Dataset [135]	2015	Derhab *et al*. [64]	SDN/NFV	Evaluate the effectiveness and efficiency of random subspace learning-based IDS
AWID dataset [136]	2015	Zhou *et al*. [65]	SDN/NFV	Evaluate an IDS based on feature selection and ensemble classifier techniques
		Lopez-Martin *et al*. [110]	IoT devices	Train and validate an IDS model based on deep reinforcement learning
CSE-CIC-IDS2018 dataset [134]	2018	Li *et al*. [98]	Autonomous tractors	Evaluate an auto-encoder IDS
Bot-IoT dataset [137]	2019	Ferrag *et al*. [99]	IoT devices	Evaluate a rules and decision tree-based intrusion detection system
TON_IOT dataset [138]	2021	Alsaedi *et al*. [140]	IoT and IIoT devices	Evaluate multiple ML methods and a DL model
InSDN dataset [139]	2020	Said Elsayed *et al*. [141]	SDN/NFV	Evaluate an LSTM-based autoencoder IDS

14) InSDN Dataset: This dataset was introduced by [139] to overcome the limitations of existing datasets in the context of SDNs. It includes more than 80 features in CSV format, with a total instances of 343 939 for both normal and attack traffic.

The InSDN dataset incorporates diverse classes of attacks such as DoS, DDoS, Web, Password-Guessing and Botnets. Furthermore, the normal traffic in InSDN includes various types of famous application services such as HTTPS, HTTP and DNS.

C. Building Process

A great deal of research has been conducted to develop intelligent intrusion detection techniques to improve the security of networks [65], [144]–[146]. The main steps for each IDS building process are 1) data collection, 2) data pre-processing, 3) classifier training, and 4) attack recognition. We provide a brief description of each step, as well as some of the latest techniques that can be implemented in Agriculture 4.0, as illustrated in Fig. 12.

1) Data Collection: Information gathering represents the first, and critical, step in intrusion detection. The type of source of the data and the point at which the data is collected are two key factors in the design and performance of an IDS. [144]. Sadikin *et al.* [145] provided a new, efficient, and reliable method for large-scale IDS data collection, which is applicable for Agriculture 4.0. The authors used Zigbee Diagnostic Reports to ensure that IDS data collection can be done reliably and efficiently in a resource-limited Zigbee IoT environment.

2) Data Pre-Processing: Once the data is obtained at the data collection stage, they are first processed in order to generate basic features [144]. The feature selection technique, used as a pre-processing step in ML algorithms, attempts to reduce the complexity of the calculations by removing unnecessary features while preserving or even improving the performance of the IDS [65]. The trained classifier demands that each record in the input data is expressed as a real number vector. Therefore, every symbolic characteristic in the dataset must be first converted to a numeric value, a technique called data transferring [144]. Data normalization refers to the process of scaling the value of each attribute across a well-proportioned range, thereby removing the bias towards characteristics with larger values from the dataset, which can increase significantly the accuracy of the classifying algorithm [144]. Khan *et al*. [147] proposed an approach called HML-IDS to address the challenge of building an intrusion detection framework from unbalanced intrusion datasets, specifically designed for the industrial control system (ICS), that is suitable for application in Agriculture 4.0. The approach includes a feature extraction technique derived from data normalization with data feature retrieval (DFR) and used a modified nearest-neighbor rule algorithm to balance the dataset, which enhanced the accuracy of classifiers. Experimental results obtained from a large-scale real dataset created using a SCADA system showed a 97% accuracy rate.

3) Classifier Training: After selecting the ideal subset of features, it is then brought into the classifier training stage [144]. To enhance the accuracy of IDS, Zhou *et al*. [65] trained three separate classifiers as basic learners using C4.5, random forest (RF) and forest by penalizing attributes (Forest PA) algorithms, then construct an ensemble classifier from them, which can be adapted to Agriculture 4.0. The experimental results showed good results with a classification accuracy of 99.81%, 99.8% DR, and 0.08% FAR with a subset of 10 characteristics for the NSL-KDD dataset. In addition, the results obtained for the AWID showed an accuracy of 99.52% with 0.15% FAR using a subset of only 8 characteristics.

4) Attack Recognition: Having completed all the above steps, it is possible to identify both normal and intrusion traffic via the trained and registered classifier. Afterward, the test data is passed to the trained and saved model for intrusion detection [144]. Ren *et al*. [146] proposed a data optimization approach to build IDS, referred to as DO_IDS, which has two main parts: data sampling and feature selection. In data sampling, the authors used iForest to sample the data while the combination of the genetic algorithm (GA) and RF is applied to optimize the sampling ratio. In feature selection, the GA and RF combination is used again but this time to select the ideal subset of characteristics. DO_IDS achieved better results than the RF classifier, in particular for the detection of anomalies, such as DoS, analysis, backdoor, and worms. Nevertheless, some enhancements can be focused on, such as the time cost associated with the data optimization step and the support for online processing.

D. IDS Potentials for Agriculture 4.0

The integration of intelligent agricultural systems makes both intelligent objects more effective and agricultural production more efficient [27]. Nevertheless, these systems are susceptible to a variety of security attacks, which can cause considerable damage to agricultural services and applications such as smart crop/livestock monitoring, food supply chain traceability, unmanned aerial vehicle (UAV), and unmanned ground vehicle (UGV) autonomous tasks. An IDS implemented for the Agriculture 4.0 environment is expected to deliver real-time packet analysis and feedback, support different network layers with different protocol stacks, and support different technologies such as IoT, Fog/Cloud computing, Blockchain, and SDN/NFV. Furthermore, it is expected to operate within tight constraints of limited processing conditions, high response speed, and high data volume processing. Table VI presents some of the benefits of IDS for Agriculture 4.0, in terms of technologies and applications applied in Agriculture 4.0, the scenarios of either the absence or presence of IDS, and the type of IDS appropriate for each scenario.

E. Software Implementations

Table VII highlights key features offered by the most popular software packages for implementing deep learning in IDS-secured Agriculture 4.0. Supported techniques include common deep learning architectures such as convolutional neural networks (CNN), recurrent neural networks (RNN), restricted Boltzmann machine (RBM), and deep belief network (DBN). Parallel processing techniques are used to boost the performance through GPU acceleration, where some of the well known techniques are CUDA, OpenMP, and OpenCL. Pre-trained models indicate whether the framework accepts already trained models as a starting point. Cloud Support states whether the framework enables cloud-based

TABLE VI
IDS'S BENEFITS TO AGRICULTURE 4.0

Technology	Application	Without an IDS	With an IDS	Suitable IDS type
Internet of Things	Smart crop/livestock monitoring	- Exploiting IoT devices - Protocol-based attacks - Crop/livestock losses - Surveillance gaps -Sensor data obstruction	Intelligent farms equipped with IDS ensures the security of IoT systems, by mitigating attacks such as DoS, RPL, and sinkhole. This helps to avoid agricultural production losses.	- A data mining-based IDS extracts knowledge from massive amounts of data and automatically generates traffic-dependent models based on traffic patterns. It is ideal for an unlimited, continuous and rapidly growing online data flow.
	Smart water management	- Malicious commands injection - Insufficient or over irrigation - Water loss	IoT-based intelligent irrigation systems could be protected by IDS against false control injection, DDoS attacks, which prevent crop and soil losses.	- The design of an IDS based on this approach consists of a rule learning step, a clustering step, a classification step, and a regression step [148].
	Disease management	- Falsification of health-related collected data - Sensible data theft - Controlling health-related devices	IDS can protect IoT-based healthcare appliance information found in livestock and crops, and mitigate cyber attacks such as data theft, and integrity issues.	- Designing a IDS-based payload model is another option for this type of scenario, where it is essentially based on port or host-specific packet traffic for a given port or host for a given agricultural application.
	Agrochemicals applications	- Malicious commands injection in agrochemicals sprayers - Irregular applications of fertilizers and diverse pesticides - Crop and soil loss	IDS's attacks detection together with fast alarm generation features provide protection against malicious sprayer control, and rule-based anomalies.	- For a signature-based IDS, an attack pattern is determined based on model matching of the payload model with specific characteristics. On the other hand, an anomaly-based IDS produces a model based on packet payloads that describe normal behavior.
Fog/Cloud computing	Cloud-based data storage	- Agricultural IoT platform cyber attacks such as SQL injection, XSS, DDoS - Critical data losses or falsification - Authentication attacks - Malware injection attacks	IDS-secured systems, whether network-based or host-based, could prevent cyber attacks on intelligent agriculture platforms hosted in the cloud, and prevent unauthorized access to information.	- IDSs based on machine learning have two stages: 1) The learning step involves mathematical algorithms that rely on normal data as the reference input to learn the characteristics of the environment. 2) The detection step uses these characteristics to detect and classify attacks [148].
	Fog-based data processing	- Fog nodes traffic jamming - Delayed or malicious field decisions - Fog nodes deauthentication attacks - Financial and resource losses - Malicious fog devices	Feature selection as well as classification-based intrusion detection techniques, such as fuzzy techniques, neural networks and genetic algorithms, are used for the protection of network security and the improvement of the quality of service (QoS).	- There are two major types of machine learning techniques: 1) Supervised learning, where the characteristics of the training data set are fed into the learning stage creating a classification pattern, making it possible to classify new, unnoticed incidents. 2) Unsupervised learning that relies on data features without using aggregated training data.
Blockchain	Food supply chain traceability	- Alter the manufacturing process of agriculturals product through the installation of malwares - DDoS the consortium Blockchain	Merging the Blockchain-based supply management system with the IDS could provide a more secure way to ensure traceability of agricultural products and manufacturing processes.	- The IDS, based on a rules model, establishes rules for the computing environment, extracted from data traffic patterns, and recognizes as an attack any abnormal data traffic that violates these rules [148].
	Food safety and quality control	- False data injection about agricultural products leading to commercial fraud - Exploit vulnerable Quality Control (QC) systems	Ensuring the integrity of data sources using IDS could mitigate against fraudulent market actions such as counterfeiting.	- The process of rule creation relies on the historical behavior of the system. Therefore, it is necessary to continuously monitor the system over a long period of time to prevent too high a false positive rate.
SDN/NFV	Agricultural IoT-based network management	- OpenFlow protocol attacks - DDoS or Hijack SDN-enabled switches or controllers - Delayed or interrupted autonomous agricultural field tasks under network failure	IDS collects statistical flow information from the OpenFlow SDN-enabled switches and evaluates traffic information through the extraction and combination of a set of features, to mitigate cyber attacks.	- The protocol-based IDS approach involves protocol surveillance in multiple layers within the network infrastructure. Using this approach, the IDS identifies anomalies related to a specific protocol which is not present within the normal model [148].
Agricultural robotics	UGV/UAV autonomus tasks	- Malicious commands injection - Network traffic jamming - Agricultural machinery and production damages - GPS spoofing attack - Compromised surveillance	Smart farms equipped with anomaly-based IDS creates a model of normal farm machinery behavior, which is continuously updated, using data from normal use, and then applying this model to detect any deviations from normal behavior.	- The signal processing-based IDS relies on signal processing methods for traffic analysis. Therefore, by capturing normal data traffic statistics as well as data frequency patterns over time, the IDS creates a normal model, with any variation from this model considered an anomaly [148].

services to accelerate the training process.

V. FUTURE DIRECTIONS

As shown in Fig. 13, to complete our study, we outline both open challenges and future research opportunities that could improve the capabilities and effectiveness of machine learning and deep learning techniques for cyber security in Agriculture 4.0, summarized in the following suggestions:

A. Audio Recognition and Computer Vision for Cyber Security Intrusion Detection

The results in our study show that deep learning techniques can provide better performance in cyber security intrusion detection for Agriculture 4.0 compared to traditional machine

TABLE VII
DEEP LEARNING FRAMEWORKS FOR IDS-SECURED AGRICULTURE 4.0

Framework	Supported platforms				Programming languages		Supported techniques				PM	PP	CS	License
	Linux	MacOS	Windows	Mobile	Core	Interface	CNN	RNN	RBM	DBN				
TensorFlow [149]	√	√	√	√	C++, Python	Python, C/C++, JAVA	√	√	√	√	√	√	√	Apache 2.0
CNTK [150]	√	×	√	×	C++	C++, Cmd	√	√	×	×	√	√	√	MIT
SINGA [151]	√	√	√	×	C++	Python, C++	√	√	√	√	√	√	×	Apache 2.0
Caffe [152]	√	√	√	×	C++	Python,C++, MATLAB	√	√	×	×	√	√	×	BSD
DL4j [153]	√	√	√	√	C++, JAVA	Python, JAVA, Scala	√	√	√	√	√	√	√	Apache 2.0
PyTorch [154]	√	√	√	×	Python, C/C++	Python, C++	√	√	√	√	√	√	√	BSD
Keras [155]	√	√	√	×	Python	Python, R	√	√	×	×	√	√	×	MIT
MXNet [156]	√	√	√	√	C++	C++, Python, R, GO	√	√	√	√	√	√	√	Apache 2.0
Theano [157]	√	√	√	√	Python	Python	√	√	√	√	√	√	×	BSD

Supported (√); Unsupported (×); Parallel processing (PP); Pre-trained models (PM); Cloud support (CS);

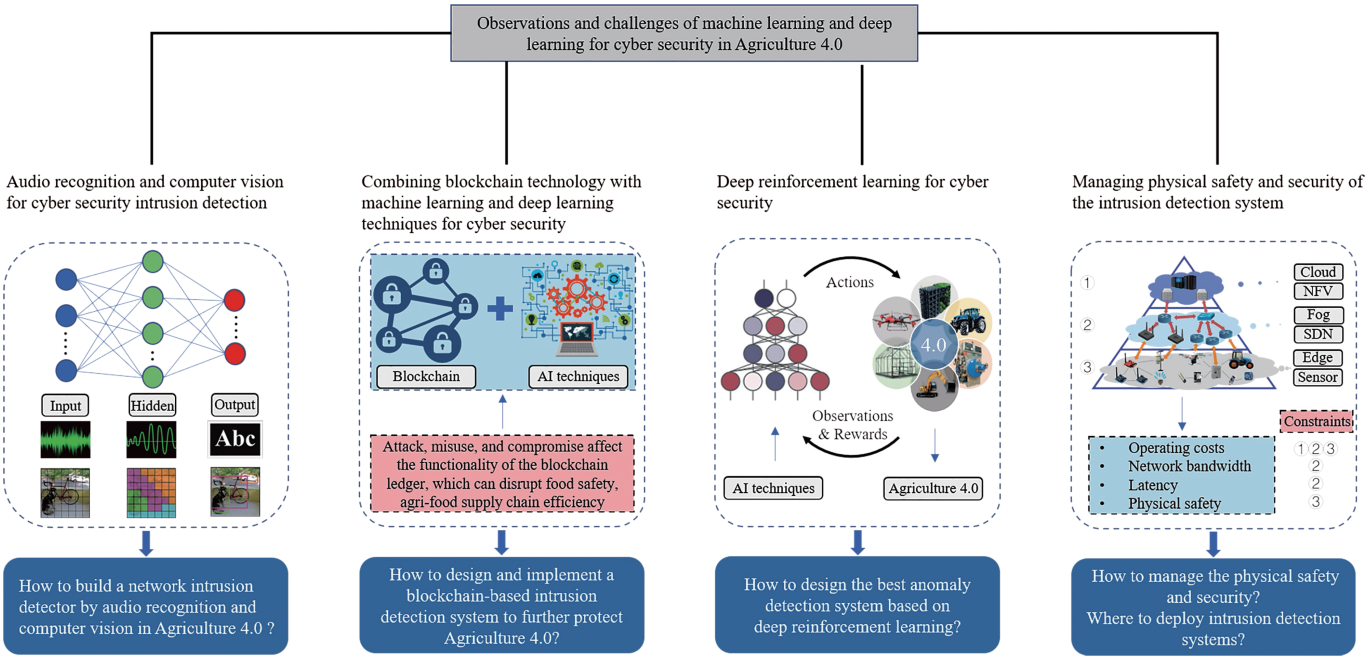

Fig. 13. Observations and challenges of machine learning and deep learning for cyber security in Agriculture 4.0.

learning techniques such as decision trees, random forests, naive Bayes, and logistic regression. Therefore, many characteristics should be taken into account when developing more efficient cyber security intrusion detection for Agriculture 4.0 such as the number of features in flow-based IDS datasets. To successfully train deep learning techniques, the number of features in flow-based IDS datasets should be adequate. Potential future research directions in this topic could be related to developing a mechanism to transform a group of network traffic flows into large encoded data, signal, or image, such as the use of Internet of multimedia things (IoMT) approaches [158]. This mechanism will train deep learning techniques well, which enhances performance of cyber security intrusion detection in Agriculture 4.0. The use of audio recognition and computer vision to build a network intrusion detector based on deep learning techniques is one significant research challenge.

B. Combining Blockchain Technology With Machine Learning and Deep Learning Techniques for Cyber Security

Blockchain is a technology for the storage and transmission of information, transparent, secure, and operating without a central control unit [159], [160]. The Blockchain can be applied for Agriculture 4.0 for improving the speed of transactions, safety, and reliability, and to reduce existing transaction or centralization costs in traditional agriculture systems [87]. However, several challenges remain for the practical realization of Blockchain in Agriculture 4.0 due to cyber security threats. For example, an adversary can launch attacks, misuses, and compromises to affect the functionality

of the Blockchain ledger, which can disrupt food safety, agri-food supply chain efficiency, and agricultural productivity. Hence, the combination of a blockchain-based scheme with an IDS must be designed and implemented to further protect Agriculture 4.0 from cyber attacks.

C. Deep Reinforcement Learning for Cyber Security

Machine learning and deep learning techniques are useful tools to detect anomalies in Agriculture 4.0. The successful implementation of these techniques depends on the following characteristics: 1) neural network opacity, 2) identification of data anomalies and errors in the dataset, 3) achieving the lowest false alarm rate, and 4) providing the highest possible accuracy. However, reinforcement learning has had success in scaling to decision-making problems [161], [162]. To solve complex problems in cyber security, the combination of deep learning and reinforcement learning, called deep reinforcement learning, can be used to detect anomalies before they do any damage in Agriculture 4.0 with the lowest false alarm rate and the highest possible accuracy. Hence, cyber security intrusion detection based on deep reinforcement learning should be carefully designed to find the best anomaly detection system.

D. Managing Physical Safety and Security of the IDS

Firewalls and anti-malware software are not sufficient to provide security in Agriculture 4.0 against cyber attacks. The use of an IDS is vital to defend against intrusion, which operates in conjunction with authorization access control, privacy, and authentication tools. Making it possible to monitor resource access events more accurately, to make sure that authorized users are granted access to information resources under certain specified conditions [163]. Since Agriculture 4.0 uses many emerging technologies (e.g., Cloud computing, Fog/Edge computing, SDN/NFV, Drones, etc.), the successful location and implementation of an IDS depends on the following constraints: 1) Lowering operating costs of the implementation in each level layer (i.e., Cloud layer, SDN layer, Edge layer, etc.), 2) Network bandwidth constraints, 3) Latency constraints, and 4) Managing physical safety and security. A possible research direction in this topic could be related to managing the physical safety and security of the IDS in Agriculture 4.0. In particular, ultrasonic and infrasound may be a potential physical attack source if they have an impact on the health of livestock or plants. In addition, the issue of deciding where IDSs should be deployed with Agriculture 4.0 remains a very challenging one to tackle.

E. Scalability in the Age of Next-Generation Networks

Next-generation networks are set to power the future of an intelligent, ubiquitous, connected, and data-rich Internet of everything (IoE), and transform the way wireless technologies evolve from "connected objects" to "connected intelligence", with extremely higher throughput. To improve performance, in many cases scientists invest more computational resources into implementation. But still, in the world of AI, no successful large-scale implementations of machine learning exists [164]. Given the fact that the Agriculture 4.0 data is heterogeneous, complex, and massive, a possible research direction would involve the development of a scalable IDS for Agriculture 4.0, with zero false positives and real-time detection.

F. Protections Against Data Poisoning Attacks

One of the most significant challenges of machine learning algorithms is their vulnerability to poisoning attacks, especially in cyber security-related applications. It is possible for an adversary to inject malicious points into the training dataset to affect the learning process [165]. This type of attack can render the IDS useless, which opens the door to many different attacks, potentially resulting in a disastrous impact on the Agriculture 4.0 industry. A future research direction should address this issue with non-computationally intensive methods to determine which regions of the underlying data collection distribution are potentially more vulnerable to data poisoning.

G. Moving Forward With Quantum Enhanced Machine Learning for IDS

Quantum computers are designed to achieve quantum supremacy by leveraging quantum physics to accomplish computational tasks that exceed the capabilities of the most powerful conventional computers [166]. A possible research direction would be the study of quantum-assisted machine learning (QAML) in IDS design and implementation for Agriculture 4.0.

VI. Conclusion

In this paper, we reviewed and analyzed IDS for cyber security in Agriculture 4.0. First, we presented the cyber security threats and the several evaluation metrics employed in evaluating the performance of an IDS for Agriculture 4.0. Next, we evaluated IDSs in relation to emerging technologies. In addition, we provided a comprehensive classification of IDSs in every emerging technology. Then, we presented the public datasets and implementation frameworks applicable to the IDS performance evaluation for Agriculture 4.0. Finally, we highlighted the challenges and future research directions in cyber security intrusion detection for Agriculture 4.0.

References

[1] Y. Liu, X. Y. Ma, L. Shu, G. P. Hancke, and A. M. Abu-Mahfouz, "From industry 4.0 to agriculture 4.0: Current status, enabling technologies, and research challenges," *IEEE Trans. Ind. Inf.*, vol. 17, no. 6, pp. 4322–4334, Jun. 2021.

[2] G. Aceto, V. Persico, and A. Pescapé, "A survey on information and communication technologies for industry 4.0: State-of-the-art, taxonomies, perspectives, and challenges," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 4, pp. 3467–3501, Aug. 2019.

[3] Industry 4.0 and cybersecurity: Managing risk in an age of connected production [Online]. Available: https://www2.deloitte.com/content/dam/insights/us/articles/3749_Industry4-0_cybersecurity/DUP_Industry4-0_cybersecurity.pdf, Accessed on: May 30, 2020.

[4] K. Huang, L. Shu, K. L. Li, F. Yang, G. J. Han, X. C. Wang, and S. Pearson, "Photovoltaic agricultural internet of things towards realizing

the next generation of smart farming," *IEEE Access*, vol. 8, pp. 76300–76312, Apr. 2020.

[5] O. Friha, M. A. Ferrag, L. Shu, and M. Nafa, "A robust security framework based on blockchain and SDN for fog computing enabled agricultural internet of things," in *Proc. Int. Conf. Internet Things and Intelligent Applications*, Zhenjiang, China, 2020, pp. 1−5.

[6] A. Tewari and B. Gupta, "Security, privacy and trust of different layers in internet-of-things (IoTs) framework," *Future Gener. Comput. Syst.*, vol. 108, pp. 909–920, Jul. 2020.

[7] W. J. Zhu, M. L. Deng, and Q. L. Zhou, "An intrusion detection algorithm for wireless networks based on ASDL," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 92–107, Jan. 2018.

[8] M. Agarwal, S. Purwar, S. Biswas, and S. Nandi, "Intrusion detection system for PS-poll DoS attack in 802.11 networks using real time discrete event system," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 4, pp. 792–808, 2017.

[9] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *J. Inf. Secur. Appl.*, vol. 50, p. 102419, Feb. 2020.

[10] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 2, pp. 1153–1176, 2016.

[11] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Cluster Comput.*, vol. 22, no. 1, pp. 949–961, Jan. 2019.

[12] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. J. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 3, pp. 1646–1685, Apr. 2020.

[13] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 1, pp. 686–728, Feb. 2019.

[14] K. A. P. da Costa, J. P. Papa, C. O. Lisboa, R. Munoz, and V. H. C. de Albuquerque, "Internet of things: A survey on machine learning-based intrusion detection approaches," *Comput. Netw.*, vol. 151, pp. 147–157, Mar. 2019.

[15] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 3, pp. 2671–2701, Jan. 2019.

[16] H. Y. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," *Appl. Sci.*, vol. 9, no. 20, p. 4396, Oct. 2019. DOI: 10.3390/app9204396.

[17] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 2, pp. 493–501, Mar. 2019.

[18] A. Ahmad, E. Harjula, M. Ylianttila, and I. Ahmad, "Evaluation of machine learning techniques for security in SDN," in *Proc. IEEE Globecom Workshops*, Taiwan, China, 2020, pp. 1−6.

[19] Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 1, p. e4150, Jan. 2021.

[20] M. Mohammadi, T. A. Rashid, S. H. T. Karim, A. H. M. Aldalwie, Q. T. Tho, M. Bidaki, A. M. Rahmani, and M. Hosseinzadeh, "A comprehensive survey and taxonomy of the SVM-based intrusion detection systems," *J. Netw. Comput. Appl.*, vol. 178, p. 102983, Mar. 2021. DOI: 10.1016/j.jnca.2021.102983.

[21] M. A. Ferrag, M. Babaghayou, and M. A. Yazici, "Cyber security for fog-based smart grid SCADA systems: Solutions and challenges," *J. Inf. Secur. Appl.*, vol. 52, p. 102500, Jun. 2020.

[22] P. P. Ray, "Internet of things for smart agriculture: Technologies, practices and future direction," *J. Amb. Intel. Smart Environ.*, vol. 9, no. 4, pp. 395–420, Jun. 2017.

[23] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, Apr. 2018.

[24] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. H. D. N. Hindia, "An overview of internet of things (IoT) and data analytics in agriculture: Benefits and challenges," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3758–3773, Oct. 2018.

[25] A. Khanna and S. Kaur, "Evolution of internet of things (IoT) and its significant impact in the field of precision agriculture," *Comput. Electron. Agric.*, vol. 157, pp. 218–231, Feb. 2019.

[26] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Comput. Electron. Agric.*, vol. 170, p. 105256, Mar. 2020. DOI: 10.1016/j.compag.2020.105256.

[27] X. Yang, L. Shu, J. N. Chen, M. A. Ferrag, J. Wu, E. Nurellari, and K. Huang, "A survey on smart agriculture: Development modes, technologies, and security and privacy challenges," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 2, pp. 273–302, Feb. 2021.

[28] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, and X. C. Wang, "Internet of things for the future of smart agriculture: A comprehensive survey of emerging technologies," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 4, pp. 718–752, Apr. 2021.

[29] K. Demestichas, N. Peppes, and T. Alexakis, "Survey on security threats in agricultural IoT and smart farming," *Sensors*, vol. 20, no. 22, p. 6458, Nov. 2020. DOI: 10.3390/s20226458.

[30] Threats to precision agriculture [Online]. Available: https://www.dhs.gov/sites/default/files/publications/2018%20AEP_Threats_to_Precision_Agriculture.pdf, Accessed on: Mar. 3, 2021.

[31] N. Munaiah, A. Meneely, R. Wilson, and B. Short, "Are intrusion detection studies evaluated consistently? A systematic literature review," 2016.

[32] A. Milenkoski, M. Vieira, S. Kounev, A. Avritzer, and B. D. Payne, "Evaluating computer intrusion detection systems: A survey of common practices," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 1–41, Sept. 2015.

[33] J. Opitz and S. Burst, "Macro F1 and macro F1," arXiv preprint arXiv: 1911.03347, Nov. 2019.

[34] H. Narasimhan, H. G. Ramaswamy, A. Saha, and S. Agarwal, "Consistent multiclass algorithms for complex performance measures," in *Proc. 32nd Int. Conf. Machine Learning*, Lille, France, 2015, pp. 2398−2407.

[35] K. S. Gill, S. Saxena, and A. Sharma, "GTM-CSec: Game theoretic model for cloud security based on ids and honeypot," *Comput. Secur.*, vol. 92, p. 101732, May 2020. DOI: 10.1016/j.cose.2020.101732.

[36] M. Rabbani, Y. L. Wang, R. Khoshkangini, H. Jelodar, R. X. Zhao, and P. Hu, "A hybrid machine learning approach for malicious behaviour detection and recognition in cloud computing," *J. Netw. Comput. Appl.*, vol. 151, p. 102507, Feb. 2020. DOI: 10.1016/j.jnca.2019.102507.

[37] G. S. Kushwah and V. Ranga, "Voting extreme learning machine based distributed denial of service attack detection in cloud computing," *J. Inf. Secur. Appl.*, vol. 53, p. 102532, Aug. 2020.

[38] A. Aldribi, I. Traoré, B. Moa, and O. Nwamuo, "Hypervisor-based cloud intrusion detection through online multivariate statistical change tracking," *Comput. Secur.*, vol. 88, p. 101646, Jan. 2020. DOI: 10.1016/j.cose.2019.101646.

[39] P. Y. Zhang, M. C. Zhou, and G. Fortino, "Security and trust issues in fog computing: A survey," *Future Gener. Comput. Syst.*, vol. 88, pp. 16–27, Nov. 2018.

[40] Z. H. Tian, C. C. Luo, J. Qiu, X. J. Du, and M. Guizani, "A distributed deep learning system for web attack detection on edge devices," *IEEE Trans. Ind. Inf.*, vol. 16, no. 3, pp. 1963–1971, Mar. 2020.

[41] A. S. Almogren, "Intrusion detection in edge-of-things computing," *J. Paral. Distrib. Comput.*, vol. 137, pp. 259–265, Mar. 2020.

[42] Z. P. Jiang, K. Zhao, R. Li, J. Z. Zhao, and J. Z. Du, "PHYAlert: Identity spoofing attack detection and prevention for a wireless edge network," *J. Cloud Comput.*, vol. 9, no. 1, pp. 1–13, Jan. 2020.

[43] Z. D. Wu, J. J. Wang, L. Q. Hu, Z. Zhang, and H. Wu, "A network intrusion detection method based on semantic Re-encoding and deep learning," *J. Netw. Comput. Appl.*, vol. 164, p. 102688, Aug. 2020. DOI: 10.1016/j.jnca.2020.102688.

[44] M. Ahsan, M. Mashuri, M. H. Lee, H. Kuswanto, and D. D. Prastyo, "Robust adaptive multivariate hotelling's *T*.2 control chart based on kernel density estimation for intrusion detection system," *Expert Syst. Appl.*, vol. 145, p. 113105, May 2020. DOI: 10.1016/j.eswa.2019.113105.

[45] A. S. Qureshi, A. Khan, N. Shamim, and M. H. Durad, "Intrusion detection using deep sparse auto-encoder and self-taught learning," *Neural Comput. Appl.*, vol. 32, no. 9, pp. 3135–3147, Apr. 2020.

[46] W. Haider, N. Moustafa, M. Keshk, A. Fernandez, K. K. R. Choo, and A. Wahab, "FGMC-HADS: Fuzzy Gaussian mixture-based correntropy models for detecting zero-day attacks from Linux systems," *Comput. Secur.*, vol. 96, p. 101906, Sep. 2020. DOI: 10.1016/j.cose.2020.101906.

[47] N. Moustafa, "New generations of internet of things datasets for cybersecurity applications based machine learning: TON_IoT datasets," in *Proc. eResearch Australasia Conf.*, Brisbane, Australia, 2019, pp. 1−2.

[48] KDD cup 1999 data [Online]. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, Accessed on: May 30, 2019.

[49] W. Haider, J. Hu, J. Slay, B. P. Turnbull, and Y. Xie, "Generating realistic intrusion detection system dataset based on fuzzy qualitative modeling," *J. Netw. Comput. Appl.*, vol. 87, pp. 185–192, Jun. 2017.

[50] B. Naik, M. S. Obaidat, J. Nayak, D. Pelusi, P. Vijayakumar, and S. H. Islam, "Intelligent secure ecosystem based on metaheuristic and functional link neural network for edge of things," *IEEE Trans. Ind. Inf.*, vol. 16, no. 3, pp. 1947–1956, Mar. 2020.

[51] A. N. Toosi and M. Kahani, "A new approach to intrusion detection based on an evolutionary soft computing model using neuro-fuzzy classifiers," *Comput. Commun.*, vol. 30, no. 10, pp. 2201–2212, Jul. 2007.

[52] B. Pfahringer, "Winning the KDD99 classification cup: Bagged boosting," *ACM SIGKDD Explor. Newsl.*, vol. 1, no. 2, pp. 65–66, Jan. 2000.

[53] M. Pawlicki, M. Choraś, and R. Kozik, "Defending network intrusion detection systems against adversarial evasion attacks," *Future Gener. Comput. Syst.*, vol. 110, pp. 148–154, Sept. 2020.

[54] S. Kaur and M. Singh, "Hybrid intrusion detection and signature generation using deep recurrent neural networks," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7859–7877, Jun. 2020.

[55] S. Hosseini and B. M. H. Zade, "New hybrid method for attack detection using combination of evolutionary algorithms, SVM, and ANN," *Comput. Netw.*, vol. 173, p. 107168, May 2020. DOI: 10.1016/j.comnet.2020.107168.

[56] S. H. Teng, N. Q. Wu, H. B. Zhu, L. Y. Teng, and W. Zhang, "SVM-DT-based adaptive and collaborative intrusion detection," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 108–118, Jan. 2018.

[57] D. M. Ngo, C. Pham-Quoc, and T. N. Thinh, "Heterogeneous hardware-based network intrusion detection system with multiple approaches for SDN," *Mobile Netw. Appl.*, vol. 25, no. 3, pp. 1178–1192, Jun. 2020.

[58] M. A. B. Ahmadon, S. Yamaguchi, Z. L. Gou, and B. B. Gupta, "Detection and update method for attack behavior models in intrusion detection systems," in *Proc. 3rd Int. Conf. Computing for Sustainable Global Development*, New Delhi, India, 2016, pp. 2119−2124.

[59] Z. L. Gou, M. A. B. Ahmadon, S. Yamaguchi, and B. B. Gupta, "A petri net-based framework of intrusion detection systems," in *Proc. IEEE 4th Global Conf. Consumer Electronics*, Osaka, Japan, 2015, pp. 579−583.

[60] I. H. Abdulqadder, S. J. Zhou, D. Q. Zou, I. T. Aziz, and S. M. A. Akber, "Multi-layered intrusion detection and prevention in the SDN/NFV enabled cloud of 5G networks using AI-based defense mechanisms," *Comput. Netw.* , vol. 179, p. 107364, Oct. 2020. DOI: 10.1016/j.comnet.2020.107364.

[61] I. H. Abdulqadder, D. Q. Zou, I. T. Aziz, B. Yuan, and W. Q. Dai, "Deployment of robust security scheme in SDN based 5G network over NFV enabled cloud environment," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 2, pp. 866–877, Apr.–Jun. 2021.

[62] W. Zong, Y. W. Chow, and W. Susilo, "Interactive three-dimensional visualization of network intrusion detection data for machine learning," *Future Gener. Comput. Syst.*, vol. 102, pp. 292–306, Jan. 2020.

[63] A. Mishra, N. Gupta, and B. B. Gupta, "Defense mechanisms against DDoS attack based on entropy in SDN-cloud using POX controller," *Telecommun. Syst.*, vol. 77, no. 1, pp. 47–62, Jan. 2021.

[64] A. Derhab, M. Guerroumi, A. Gumaei, L. Maglaras, M. A. Ferrag, M. Mukherjee, and F. A. Khan, "Blockchain and random subspace learning-based IDS for SDN-enabled industrial IoT security," *Sensors*, vol. 19, no. 14, p. 3119, Jul. 2019. DOI: 10.3390/s19143119.

[65] Y. Y. Zhou, G. Cheng, S. Q. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Comput. Netw.*, vol. 174, p. 107247, Jun. 2020. DOI: 10.1016/j.comnet.2020.107247.

[66] L. Lv, W. H. Wang, Z. Y. Zhang, and X. G. Liu, "A novel intrusion detection system based on an optimal hybrid kernel extreme learning machine," *Knowl.-Based Syst.*, vol. 195, p. 105648, May 2020. DOI: 10.1016/j.knosys.2020.105648.

[67] S. Velliangiri and P. Karthikeyan, "Hybrid optimization scheme for intrusion detection using considerable feature selection," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7925–7939, Jun. 2019.

[68] P. Chellammal and K. M. P. D. Sheba, "Real-time anomaly detection using parallelized intrusion detection architecture for streaming data," *Concurr. Comput. Pract. Exper.*, vol. 32, no. 4, p. e5013, Feb. 2020.

[69] C. Khammassi and S. Krichen, "A NSGA2-LR wrapper approach for feature selection in network intrusion detection," *Comput. Netw.*, vol. 172, p. 107183, May 2020. DOI: 10.1016/j.comnet.2020.107183.

[70] S. J. Bu and S. B. Cho, "A convolutional neural-based learning classifier system for detecting database intrusion via insider attack," *Inf. Sci.*, vol. 512, pp. 123–136, Feb. 2020.

[71] Y. Q. Yang, K. F. Zheng, B. Wu, Y. X. Yang, and X. J. Wang,

"Network intrusion detection based on supervised adversarial variational auto-encoder with regularization," *IEEE Access*, vol. 8, pp. 42169–42184, Feb. 2020.

[72] K. Y. Jiang, W. Y. Wang, A. L. Wang, and H. B. Wu, "Network intrusion detection combined hybrid sampling with deep hierarchical network," *IEEE Access*, vol. 8, pp. 32464–32476, Feb. 2020.

[73] N. Guizani and A. Ghafoor, "A network function virtualization system for detecting malware in large IoT based networks," *IEEE J. Select. Areas Commun.*, vol. 38, no. 6, pp. 1218–1228, Jun. 2020.

[74] M. A. Ferrag and L. Maglaras, "DeliveryCoin: An IDS and blockchain-based delivery framework for drone-delivered services," *Computers*, vol. 8, no. 3, p. 58, Aug. 2019. DOI: 10.3390/computers 8030058.

[75] C. Ju and H. I. Son, "Multiple UAV systems for agricultural applications: Control, implementation, and evaluation," *Electronics*, vol. 7, no. 9, p. 162, Aug. 2018. DOI: 10.3390/electronics7090162.

[76] D. Albani, T. Manoni, A. Arik, D. Nardi, and V. Trianni, "Field coverage for weed mapping: Toward experiments with a UAV swarm," in *Proc. 11th EAI Int. Conf. Bio-inspired Information and Communication*, Pittsburgh, USA, 2019, pp. 132−146.

[77] U. Challita, A. Ferdowsi, M. Z. Chen, and W. Saad, "Machine learning for wireless connectivity and security of cellular-connected UAVs," *IEEE Wirel. Commun.*, vol. 26, no. 1, pp. 28–35, Feb. 2019.

[78] S. K. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *Ad Hoc Netw.*, vol. 105, p. 102177, Aug. 2020. DOI: 10.1016/j.adhoc.2020.102177.

[79] M. Wang, Y. Q. Lu, and J. C. Qin, "A dynamic MLP-based DDoS attack detection method using feature selection and feedback," *Comput. Secur.*, vol. 88, p. 101645, Jan. 2020. DOI: 10.1016/j.cose. 2019.101645.

[80] S. Sciancalepore, O. A. Ibrahim, G. Oligeri, and R. Di Pietro, "PiNcH: An effective, efficient, and robust solution to drone detection via network traffic analysis," *Comput. Netw.*, vol. 168, p. 107044, Feb. 2020. DOI: 10.1016/j.comnet.2019.107044.

[81] Q. M. Alzubi, M. Anbar, Z. N. M. Alqattan, M. A. Al-Betar, and R. Abdullah, "Intrusion detection system based on a modified binary grey wolf optimisation," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6125–6137, May 2020.

[82] W. Elmasry, A. Akbulut, and A. H. Zaim, "Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic," *Comput. Netw.*, vol. 168, p. 107042, Feb. 2020. DOI: 10.1016/j.comnet.2019.107042.

[83] M. Al Qurashi, C. M. Angelopoulos, and V. Katos, "An architecture for resilient intrusion detection in ad-hoc networks," *J. Inf. Secur. Appl.*, vol. 53, p. 102530, Aug. 2020.

[84] N. V. Abhishek, A. Tandon, T. J. Lim, and B. Sikdar, "A GLRT-based mechanism for detecting relay misbehavior in clustered IoT networks," *IEEE Trans. Inf. Foren. Secur.*, vol. 15, pp. 435–446, Aug. 2019.

[85] M. Villamizar, A. Martínez-González, O. Canévet, and J. M. Odobez, "WatchNet++: Efficient and accurate depth-based network for detecting people attacks and intrusion," *Mach. Vision Appl.*, vol. 31, no. 6, p. 41, Jun. 2020. DOI: 10.1007/s00138-020-01089-y.

[86] W. J. Li, W. Z. Meng, and M. H. Au, "Enhancing collaborative intrusion detection via disagreement-based semi-supervised learning in IoT environments," *J. Netw. Comput. Appl.*, vol. 161, p. 102631, Jul. 2020. DOI: 10.1016/j.jnca.2020.102631.

[87] M. A. Ferrag, L. Shu, X. Yang, A. Derhab, and L. Maglaras, "Security and privacy for green IoT-based agriculture: Review, blockchain solutions, and challenges," *IEEE Access*, vol. 8, pp. 32031–32053, Feb. 2020.

[88] Y. H. Wang, Z. H. Tian, Y. B. Sun, X. J. Du, and N. Guizani, "Locjury: An IBN-based location privacy preserving scheme for IoCV," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5028–5037, Aug. 2021.

[89] H. M. Song, J. Woo, and H. K. Kim, "In-vehicle network intrusion detection using deep convolutional neural network," *Vehicular Commun.*, vol. 21, p. 100198, Jan. 2020. DOI: 10.1016/j.vehcom.2019. 100198.

[90] F. van Wyk, Y. Y. Wang, A. Khojandi, and N. Masoud, "Real-time sensor anomaly detection and identification in automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1264–1276, Mar. 2020.

[91] C. Ieracitano, A. Adeel, F. C. Morabito, and A. Hussain, "A novel statistical analysis and autoencoder driven intelligent intrusion detection approach," *Neurocomputing*, vol. 387, pp. 51–62, Apr. 2020.

[92] J. M. Vidal, M. A. S. Monge, and S. M. M. Monterrubio, "EsPADA: Enhanced payload analyzer for malware detection robust against adversarial threats," *Future Gener. Comput. Syst.*, vol. 104, pp. 159–173, Mar. 2020.

[93] K. Vieira, F. L. Koch, J. B. M. Sobral, C. B. Westphall, and J. L. de Souza Leão, "Autonomic intrusion detection and response using big data," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1984–1991, Jun. 2020.

[94] R. B. Benisha and S. R. Ratna, "Detection of data integrity attacks by constructing an effective intrusion detection system," *J. Ambient Intell. Human. Comput.*, vol. 11, no. 11, pp. 5233–5244, Mar. 2020.

[95] H. P. Zhang, L. L. Huang, C. Q. Wu, and Z. B. Li, "An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," *Comput. Netw.*, vol. 177, p. 107315, Aug. 2020. DOI: 10.1016/j.comnet.2020.107315.

[96] C. Liu, J. Yang, and J. Q. Wu, "Web intrusion detection system combined with feature analysis and SVM optimization," *EURASIP J. Wirel. Commun. Netw.*, vol. 2020, no. 1, p. 33, Feb. 2020. DOI: 10.1186/s13638-019-1591-1.

[97] M. Zhou, L. S. Han, H. W. Lu, and C. Fu, "Distributed collaborative intrusion detection system for vehicular Ad Hoc networks based on invariant," *Comput. Netw.*, vol. 172, p. 107174, May 2020. DOI: 10.1016/j.comnet.2020.107174.

[98] X. K. Li, W. Chen, Q. R. Zhang, and L. F. Wu, "Building auto-encoder intrusion detection system based on random forest feature selection," *Comput. Secur.*, vol. 95, p. 101851, Aug. 2020. DOI: 10.1016/j.cose. 2020.101851.

[99] M. A. Ferrag, L. Maglaras, A. Ahmim, M. Derdour, and H. Janicke, "RDTIDS: Rules and decision tree-based intrusion detection system for internet-of-things networks," *Future Internet*, vol. 12, no. 3, p. 44, Mar. 2020. DOI: 10.3390/fi12030044.

[100] How to choose an activation function for deep learning [Online]. https://wmk-it.net/technology/how-to-choose-an-activation-function-for-deep-learning-1627113168, Accessed on: Mar. 13, 2021.

[101] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi, and A. Razaque, "Deep recurrent neural network for IoT intrusion detection system," *Simul. Modell. Pract. Theory*, vol. 101, p. 102031, May 2020. DOI: 10.1016/j.simpat.2019.102031.

[102] D. M. Li, L. B. Deng, B. B. Gupta, H. X. Wang, and C. Choi, "A novel CNN based security guaranteed image watermarking generation scenario for smart city applications," *Inf. Sci.*, vol. 479, pp. 432–447, Apr. 2019.

[103] S. A. Aljawarneh and R. Vangipuram, "GARUDA: Gaussian

dissimilarity measure for feature representation and anomaly detection in internet of things," *J. Supercomput.*, vol. 76, no. 6, pp. 4376–4413, Jun. 2020.

[104] F. Jiang, Y. S. Fu, B. B. Gupta, Y. S. Liang, S. Rho, F. Lou, F. Z. Meng, and Z. H. Tian, "Deep learning based multi-channel intelligent attack detection for data security," *IEEE Trans. Sustain. Comput.*, vol. 5, no. 2, pp. 204–212, Apr.–Jun. 2020.

[105] W. Wang, Y. Y. Shang, Y. Z. He, Y. D. Li, and J. Q. Liu, "Botmark: Automated botnet detection with hybrid analysis of flow-based and graph-based traffic behaviors," *Inf. Sci.*, vol. 511, pp. 284–296, Feb. 2020.

[106] M. Shafiq, Z. H. Tian, A. K. Bashir, X. J. Du, and M. Guizani, "CorrAUC: A malicious bot-IoT traffic detection method in IoT network using machine-learning techniques," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3242–3254, Mar. 2020.

[107] A. Al Shorman, H. Faris, and I. Aljarah, "Unsupervised intelligent system based on one class support vector machine and grey wolf optimization for IoT botnet detection," *J. Ambient Intell. Human. Comput.*, vol. 11, no. 7, pp. 2809–2825, Jul. 2019.

[108] M. M. Hassan, A. Gumaei, A. Alsanad, M. Alrubaian, and G. Fortino, "A hybrid deep learning model for efficient intrusion detection in big data environment," *Inf. Sci.*, vol. 513, pp. 386–396, Mar. 2020.

[109] S. Murali and A. Jamalipour, "A lightweight intrusion detection for sybil attack under mobile RPL in the internet of things," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 379–388, Jan. 2020.

[110] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, "Application of deep reinforcement learning to intrusion detection for supervised problems," *Expert Syst. Appl.*, vol. 141, p. 112963, Mar. 2020. DOI: 10.1016/j.eswa.2019.112963.

[111] N. Kumar, V. Poonia, B. B. Gupta, and M. K. Goyal, "A novel framework for risk assessment and resilience of critical infrastructure towards climate change," *Technol. Forecast. Soc. Change*, vol. 165, p. 120532, Apr. 2021. DOI: 10.1016/j.techfore.2020.120532.

[112] W. Liang, K. C. Li, J. Long, X. Y. Kui, and A. Y. Zomaya, "An industrial network intrusion detection algorithm based on multifeature data clustering optimization model," *IEEE Trans. Ind. Inf.*, vol. 16, no. 3, pp. 2063–2071, Mar. 2020.

[113] S. T. Park, G. Z. Li, and J. C. Hong, "A study on smart factory-based ambient intelligence context-aware intrusion detection system using machine learning," *J. Ambient Intell. Human. Comput.*, vol. 11, no. 4, pp. 1405–1412, Apr. 2020.

[114] F. Farivar, M. S. Haghighi, A. Jolfaei, and M. Alazab, "Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial IoT," *IEEE Trans. Ind. Inf.*, vol. 16, no. 4, pp. 2716–2725, Apr. 2020.

[115] J. P. Liu, W. X. Zhang, T. Y. Ma, Z. H. Tang, Y. F. Xie, W. H. Gui, and J. P. Niyoyita, "Toward security monitoring of industrial cyber-physical systems via hierarchically distributed intrusion detection," *Expert Syst. Appl.*, vol. 158, p. 113578, Nov. 2020. DOI: 10.1016/j.eswa.2020.113578.

[116] Y. Hu, H. Li, T. H. Luan, A. Yang, L. M. Sun, Z. L. Wang, and R. Wang, "Detecting stealthy attacks on industrial control systems using a permutation entropy-based method," *Future Gener. Comput. Syst.*, vol. 108, pp. 1230–1240, Jul. 2020.

[117] K. L. Miao, X. F. Shi, and W. A. Zhang, "Attack signal estimation for intrusion detection in industrial control system," *Comput. Secur.*, vol. 96, p. 101926, Sep. 2020. DOI: 10.1016/j.cose.2020.101926.

[118] M. N. Kurt, O. Ogundijo, C. Li, and X. D. Wang, "Online cyber-attack detection in smart grid: A reinforcement learning approach," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5174–5185, Sept. 2019.

[119] A. Patel, H. Alhussian, J. M. Pedersen, B. Bounabat, J. C. Júnior, and S. Katsikas, "A nifty collaborative intrusion detection and prevention architecture for smart grid ecosystems," *Comput. Secur.*, vol. 64, pp. 92–109, Jan. 2017.

[120] L. Haghnegahdar and Y. Wang, "A whale optimization algorithm-trained artificial neural network for smart grid cyber intrusion detection," *Neural Comput. Appl.*, vol. 32, no. 13, pp. 9427–9441, Jul. 2020.

[121] M. Zhong, Y. J. Zhou, and G. Chen, "Sequential model based intrusion detection system for IoT servers using deep learning methods," *Sensors*, vol. 21, no. 4, p. 1113, Feb. 2021. DOI: 10.3390/s21041113.

[122] S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," in *Proc. IEEE 21st Int. Conf. Emerging Technologies and Factory Automation*, Berlin, Germany, 2016, pp. 1−8.

[123] W. C. Lin, S. W. Ke, and C. F. Tsai, "CANN: An intrusion detection system based on combining cluster centers and nearest neighbors," *Knowl.-Based Syst.*, vol. 78, pp. 13–21, Apr. 2015.

[124] N. Gao, L. Gao, Q. L. Gao, and H. Wang, "An intrusion detection model based on deep belief networks," in *Proc. Second Int. Conf. Advanced Cloud and Big Data*, Huangshan, China, 2014, pp. 247−252.

[125] A. Tesfahun and D. L. Bhaskari, "Intrusion detection using random forests classifier with SMOTE and feature reduction," in *Proc. Int. Conf. Cloud & Ubiquitous Computing & Emerging Technologies*, Pune, India, 2013, pp. 127−132.

[126] Y. H. Li, J. B. Xia, S. L. Zhang, J. K. Yan, X. C. Ai, and K. B. Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 424–430, Jan. 2012.

[127] M. Lezoche, J. E. Hernandez, M. del Mar Eva Alemany Díaz, H. Panetto, and J. Kacprzyk, "Agri-food 4.0: A survey of the supply chains and technologies for the future agriculture," *Comput. Ind.*, vol. 117, p. 103187, May 2020. DOI: 10.1016/j.compind.2020.103187.

[128] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proc. Military Communications and Information Systems Conf.*, Canberra, Australia, 2015, pp. 1−6.

[129] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symp. Computational Intelligence for Security and Defense Applications*, Ottawa, Canada, 2009, pp. 1−6.

[130] A. Shiravi, H. Shiravi, M. Tavallaee, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Comput. Secur.*, vol. 31, no. 3, pp. 357–374, May 2012.

[131] Isot cloud intrusion dataset [Online]. https://www.uvic.ca/engineering/ece/isot/datasets/index.php, Accessed on: May 30, 2020.

[132] HTTP DATASET CSIC 2010 [Online]. https://www.isi.csic.es/dataset/, Accessed on: May 30, 2020.

[133] Kyoto 2006 dataset [Online]. http://www.takakura.com/Kyoto_data/, Accessed on: May 30, 2020.

[134] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. 4th Int. Conf. Information Systems Security and Privac*, Funchal, Portugal, 2018, pp. 108−116.

[135] Industrial control system (ICS) cyber attack datasets [Online]. https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets, Accessed on: May 30, 2020.

[136] AWID dataset [Online]. http://icsdweb.aegean.gr/awid/, Accessed on:

May 30, 2020.

[137] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic Botnet dataset in the internet of things for network forensic analytics: Bot-IoT dataset," *Future Gener. Comput. Syst.*, vol. 100, pp. 779–796, Nov. 2019.

[138] TON_IOT datasets [Online]. https://ieee-dataport.org/documents/toniot-datasets, Mar. 3, 2021.

[139] M. S. Elsayed, N. A. Le-Khac, and A. D. Jurcut, "InSDN: A novel SDN intrusion dataset," *IEEE Access*, vol. 8, pp. 165263–165284, Sept. 2020.

[140] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems," *IEEE Access*, vol. 8, pp. 165130–165150, Sept. 2020.

[141] M. S. Elsayed, N. A. Le-Khac, S. Dev, and A. D. Jurcut, "Network anomaly detection using LSTM based autoencoder," in *Proc. 16th ACM Symp. QoS and Security for Wireless and Mobile Networks*, Alicante, Spain, 2020, pp. 37−45.

[142] S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, "Cost-based modeling for fraud and intrusion detection: Results from the JAM project," in *Proc. DARPA Information Survivability Conf. and Expo.*, Hilton Head, USA, 2000, pp. 130−144.

[143] R. P. Lippmann, D. J. Fried, I. Graf, J. W. Haines, K. R. Kendall, D. McClung, D. Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and M. A. Zissman, "Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation," in *Proc. DARPA Information Survivability Conf. and Expo.*, Los Alamitos, USA, 2000, pp. 12−26.

[144] M. A. Ambusaidi, X. J. He, P. Nanda, and Z. Y. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Trans. Comput.*, vol. 65, no. 10, pp. 2986–2998, Oct. 2016.

[145] F. Sadikin, T. van Deursen, and S. Kumar, "A ZigBee intrusion detection system for IoT using secure and efficient data collection," *Internet Things*, vol. 12, p. 100306, Dec. 2020. DOI: 10.1016/j.iot.2020.100306.

[146] J. D. Ren, J. W. Guo, W. Qian, H. Yuan, X. B. Hao, and J. J. Hu, "Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms," *Secur. Commun. Netw.*, vol. 2019, p. 7130868, Jun. 2019.

[147] I. A. Khan, D. C. Pi, Z. U. Khan, Y. Hussain, and A. Nawaz, "HML-IDS: A hybrid-multilevel anomaly prediction approach for intrusion detection in SCADA systems," *IEEE Access*, vol. 7, pp. 89507–89521, Jul. 2019.

[148] M. F. Elrawy, A. I. Awad, and H. F. A. Hamed, "Intrusion detection systems for IoT-based smart environments: A survey," *J. Cloud Comput.*, vol. 7, no. 1, p. 21, Dec. 2018. DOI: 10.1186/s13677-018-0123-6.

[149] Tensorflow [Online]. https://www.tensorflow.org, Jan. 5, 2021.

[150] The microsoft cognitive toolkit [Online]. https://docs.microsoft.com/en-us/cognitive-toolkit/, Accessed on: Jan. 5, 2021.

[151] Apache SINGA [Online]. https://singa.apache.org/, Accessed on: Jan. 5, 2021.

[152] Caffe [Online]. https://caffe.berkeleyvision.org/, Jan. 5, 2021.

[153] Eclipse deeplearning4j [Online]. https://deeplearning4j.org/, Accessed on: Jan. 5, 2021.

[154] PyTorch [Online]. https://pytorch.org/, Accessed on: Jan. 5, 2021.

[155] Keras [Online]. https://keras.io/, Accessed on: Jan. 5, 2021.

[156] Apache MXNet [Online]. https://mxnet.apache.org/, Accessed on: Jan. 5, 2021.

[157] Theano [Online]. https://pypi.org/project/Theano/, Accessed on: Jan. 5, 2021.

[158] S. AlZu'bi, B. Hawashin, M. Mujahed, Y. Jararweh, and B. B. Gupta, "An efficient employment of internet of multimedia things in smart and future agriculture," *Multim. Tools Appl.*, vol. 78, no. 20, pp. 29581–29605, Feb. 2019.

[159] X. M. Huang, D. D. Ye, R. Yu, and L. Shu, "Securing parked vehicle assisted fog computing with blockchain and optimal smart contract design," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 2, pp. 426–441, Mar. 2020.

[160] P. Y. Zhang and M. C. Zhou, "Security and trust in blockchains: Architecture, key technologies, and open issues," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 3, pp. 790–801, Jun. 2020.

[161] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.

[162] T. Liu, B. Tian, Y. F. Ai, and F. Y. Wang, "Parallel reinforcement learning-based energy efficiency improvement for a cyber-physical system," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 2, pp. 617–626, Mar. 2020.

[163] J. Qiu, Z. H. Tian, C. L. Du, Q. Zuo, S. Su, and B. X. Fang, "A survey on access control in the age of internet of things," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4682–4696, Jun. 2020.

[164] D. N. Cheng, H. P. Zhang, F. Xia, S. G. Li, Y. Q. Zhang, "The scalability for parallel machine learning training algorithm: Dataset matters," arXiv preprint arXiv: 1910.11510, 2019.

[165] L. Muñoz-González, B. Pfitzner, M. Russo, J. Carnerero-Cano, and E. C. Lupu, "Poisoning attacks with generative adversarial nets," arXiv preprint arXiv: 1906.07773, 2019.

[166] A. Gouveia and M. Correia, "Towards quantum-enhanced machine learning for network intrusion detection," in *Proc. IEEE 19th Int. Symp. Network Computing and Applications*, Cambridge, USA, 2020, pp. 1−8.

Mohamed Amine Ferrag received the bachelor degree (June, 2008), master degree (June, 2010), Ph.D. degree (June, 2014), HDR degree (April, 2019) from Badji Mokhtar-Annaba University, Algeria, all in computer science. Since October 2014, he is a Senior Lecturer at the Department of Computer Science, Guelma University, Algeria. Since July 2019, he is a Visiting Senior Researcher, NAULincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University. His research interests include wireless network security, network coding security, and applied cryptography. He is featured in Stanford University's list of the world's Top 2% Scientists for the year 2019. He has been conducting several research projects with international collaborations on these topics. He has published more than 80 papers in international journals and conferences in the above areas. Some of his research findings are published in top-cited journals, such as the *IEEE Communications Surveys and Tutorials*, *IEEE Internet of Things Journal*, *IEEE Transactions on Engineering Management*, *IEEE Access*, *Journal of Information Security and Applications* (Elsevier), *Transactions on Emerging Telecommunications Technologies* (Wiley), *Telecommunication Systems* (Springer), *International Journal of Communication Systems* (Wiley), *Sustainable Cities and Society* (Elsevier), and *Journal of Network and Computer Applications* (Elsevier). He is currently serving on various editorial positions such as Editorial Board Member in journals (Indexed SCI & Scopus) such as, *IET Networks*, *International Journal of Internet Technology and Secured Transactions* (Inderscience Publishers), *Security and Communication Networks* (Wiley), and *MDPI Journal of Sensor and Actuator Networks*. His current H-index is 20, i10-index is 29, and 1866 citations in Google Scholar Citation.

Lei Shu (Senior Member, IEEE) received the B.S. degree in computer science from South Central University for Nationalities in 2002, and the M.S. degree in computer engineering from Kyung Hee University, South Korea, in 2005, and the Ph.D. degree from the Digital Enterprise Research Institute, National University of Ireland, Ireland, in 2010. Until 2012, he was a Specially Assigned Researcher with the Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, Japan. He is currently a Distinguished Professor with Nanjing Agricultural University, and a Lincoln Professor with the University of Lincoln, U.K. He is also the Director of the NAU-Lincoln Joint Research Center of Intelligent Engineering. He has published over 400 papers in related conferences, journals, and books in the areas of sensor networks and Internet of Things. His current H-index is 62 and i10-index is 244 in Google Scholar Citation. His current research interests include wireless sensor networks and Internet of Things. He has also served as a TPC member for more than 160 conferences, such as ICDCS, DCOSS, MASS, ICC, GLOBECOM, ICCCN, WCNC, and ISCC. He was a Recipient of the 2014 Top Level Talents in Sailing Plan of Guangdong Province, China, the 2015 Outstanding Young Professor of Guangdong Province, and the GLOBECOM 2010, ICC 2013, ComManTel 2014, WICON 2016, SigTelCom 2017 Best Paper Awards, the 2017 and 2018 IEEE Systems Journal Best Paper Awards, the 2017 Journal of Network and Computer Applications Best Research Paper Award, and the Outstanding Associate Editor Award of 2017, and the 2018 IEEE ACCESS. He has also served over 60 various Co-Chair for international conferences/workshops, such as IWCMC, ICC, ISCC, ICNC, Chinacom, especially the Symposium Co-Chair for IWCMC 2012, ICC 2012, the General Co-Chair for Chinacom 2014, Qshine 2015, Collaboratecom 2017, DependSys 2018, and SCI 2019, the TPC Chair for InisCom 2015, NCCA 2015, WICON 2016, NCCA 2016, Chinacom 2017, InisCom 2017, WMNC 2017, and NCCA 2018.

Othmane Friha received the master degree in computer science from Badji Mokhtar-Annaba University, Algeria, in 2018. He is currently working toward the Ph.D. degree in the University of Badji Mokhtar-Annaba, Algeria. His current research interests include network and computer security, Internet of Things, and applied cryptography.

Xing Yang received the M.S. degree in control engineering from Nanjing University of Information Science and Technology in 2018. He is currently working toward the Ph.D. degree in the College of Engineering, Nanjing Agricultural University. His current research interests include fault diagnosis in wireless sensor networks, agricultural Internet of Things and machine learning algorithm.