DOI 10.15302/J-SSCAE-2021.02.013

Deep Learning and Industrial Internet Security: Application and Challenges

Yang Chen, Ma Ruicheng, Wang Yushi, Zhai Yanlong, Zhu Liehuang

School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China

Abstract: Industrial Internet security is crucial for strengthening the manufacturing and network sectors of China. Deep learning, owing to its strong expression ability, good adaptability, and high portability, can support the establishment of an intelligent and autonomous industrial Internet security system and method. Therefore, it is of great value to promote the integrated innovation of deep learning and industrial Internet security. In this study, we analyze the development demand for industrial Internet security from the perspective of macro industrial environment, security technology, and deep learning systems. Moreover, we summarize the application status of deep learning to industrial Internet security in terms of device, control, network, application, and data layers. The security challenges faced by deep learning applications of the industrial Internet primarily lie in model training and prediction. Furthermore, we identify key research directions, including the interpretability of deep neural networks, cost control of sample collection and calculation, imbalance of sample sets, reliability of model results, and tradeoff between availability and security. Finally, suggestions are provided: a dynamic in-depth defense system should be established in terms of overall security strategy, an application-driven and frontier exploration integrated method should be adopted to achieve breakthroughs regarding key technologies, and resource input should be raised for interdisciplinary fields to establish an industry–university–research institute joint research ecosystem.

Keywords: industrial Internet security; Internet-of-Things security; deep learning; data security

1 Introduction

The industrial Internet, deeply integrated with the latest generation of information technology and the manufacturing industry, is an emerging mode of industrial ecology and application. With the ubiquitous and reliable interconnection of human, machine, and things, it achieves the connection of all factors of production and the entire industrial and value chains and promotes the transformation of the production mode and enterprise form in the manufacturing industry. Security of the industrial Internet is a prerequisite for its high-quality development. As emphasized in the *Guidance on Strengthening the Security on the Industrial Internet* (2019), the industrial Internet security is very important and calls for an exploration of emerging technologies, such as artificial intelligence, to improve the level of its security protection.

Deep learning has a strong automatic feature extraction capability, hence providing intelligent, accurate, and advanced analysis tools for industrial Internet security in the age of big data (characterized by complex application scenarios and huge data size) [1]. Based on the original data, a series of nonlinear processing layers is used to study data representation on different levels of abstraction, and with end-to-end optimization, hidden patterns are defined and identified, and features of high nonlinearity and complexity are extracted, eliminating manual extraction of the best features from domain knowledge, and supporting data-driven industrial applications. At the

Received date: January 20, 2021**; Revised date:** March 06, 2021

Corresponding author: Zhu Liehuang, professor of School of Cyberspace Science and Technology of Beijing Institute of Technology. Major research field is network and information security. E-mail: liehuangz@bit.edu.cn

Funding program: CAE Advisory Project “Research on Development Strategy for Next-Generation Industrial Internet Security Technology” (2020-XZ-02)

Chinese version: Strategic Study of CAE 2021, 23(2): 095–103

Cited item: Yang Chen et al. Deep Learning and Industrial Internet Security: Application and Challenges. *Strategic Study of CAE*, https://doi.org/10.15302/J-SSCAE-2021.02.013

same time, it should be noted that the introduction and application of deep learning makes industrial Internet systems vulnerable to malicious attacks or illegal use (Fig. 1), which may lead to inaccurate decision-making and pose potential risks of losses in industrial manufacturing [2].

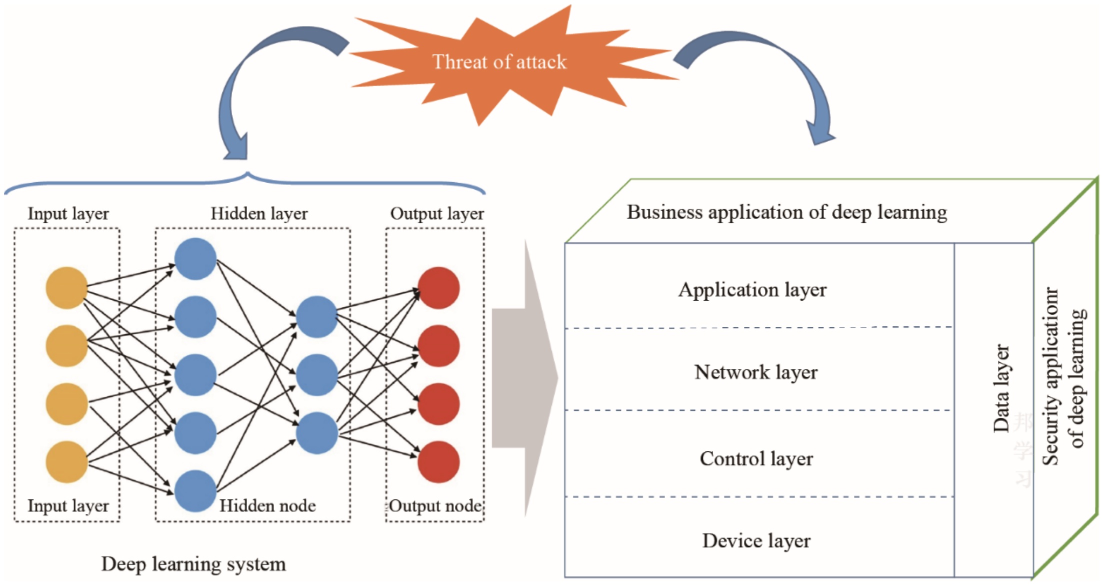

Fig. 1. Industrial Internet security is faced with the attack threat introduced by deep learning.

Although there are some studies on industrial Internet security as deep learning is being applied to the industrial Internet, these studies are relatively incomplete, and most of them fail to pay enough attention to the security of deep learning systems [3]. Therefore, this study attempts to conduct a prospective study to fill this gap. Specifically, it analyzes the security needs of industrial Internet, summarizes the specific applications of deep learning and the security challenges after the introduction of new technologies, and identifies key research directions in the field. In this way, we hope to provide strategic references for the development of industrial Internet security in China.

2 Requirement analysis of industrial Internet security

2.1 Security needs of industrial Internet

As mentioned above, industrial Internet security is the cornerstone of building China into strong manufacturing and cyber power, so it is interrelated with the high-quality development of China's economy. With manufacturing elements fully interconnected and connected to the open industrial Internet network, this interconnection brings advantages in scale and efficiency. In the meantime, it also brings potential security problems: large amounts of manufacturing resources that were originally in a closed state are now exposed to the Internet, facing a more open Internet environment and more likely to be reached and attacked maliciously by external organizations. Computing resources of the manufacturing factors are limited, and protection capability is generally weak as a result of being generated in the original closed environment, causing vulnerability to breach and illegal use. This can result in great harm, even if the destruction is only present in a single point of the entire networked collaborative industrial system considering that industrial systems generally present strict requirements for reliability, accuracy, low delay, etc. Therefore, industrial Internet applications put forward high levels of requirements for security and call for the utilization of advanced technologies such as deep learning to solve these challenges.

Technically, traditional protection measures to ensure industrial Internet security can defend against many known security threats. However, as the industrial Internet expands its field of application, the number and categories of access devices continue to increase, and emerging attacks are constantly "innovated," current traditional security defense tools and technologies tend to be less effective in dealing with new attacks. Furthermore, there is an increasing trend in the number, size, speed, and type of present-day industrial Internet attacks. All these facts indicate an urgent need for the introduction of security protection methods that are rapid, efficient, and intelligent. Because deep learning shows strong self-learning ability and excellent performance in feature discovery and automatic analysis, it can be of great value in various levels of security protection on the industrial Internet, ranging from equipment, control, application to network, data, and others, and become a feasible technical direction for protection against new forms of attacks [4].

2.2 Security requirements for deep learning systems on the industrial Internet

Deep learning technology can be widely applied to the five-tier architecture of the industrial Internet and cover all stages of the full lifecycle (Fig. 2). This can significantly reduce manual operation and improve the level of automation and production efficiency. For example, in the equipment layer, supervised deep learning is adopted to monitor the use of machines and equipment and detect the causes of failure. Combined with a voiceprint-based product quality inspection system, it can realize the automation and intelligentization of quality inspection [5]. Another example is the application layer, which adopts image recognition technology based on deep learning to achieve visual detection, sorting, positioning, etc., and to improve the efficiency and intelligence level of the assembly line. Other deep learning applications, such as demand/sales forecasts, customer profiles, and supply chain optimization, are also instrumental in assisting enterprises in making decisions [6].

Currently, there have been some studies on deep learning technology targeted for industrial Internet security, such as the intrusion detection system and data audit system. The former can achieve better detection of malicious behaviors such as scope, speed, and adaptability, whereas the latter is able to support the extraction of key information from huge amounts of industrial data and search for behaviors that threaten industrial Internet security. With the expansion and deepening of these deep learning applications, however, security issues of the deep learning system itself have become problematic as well. If these security problems are not prevented, they may pose serious hidden dangers to the industrial Internet, as this Internet requires high levels of reliability, stability, and predictability.

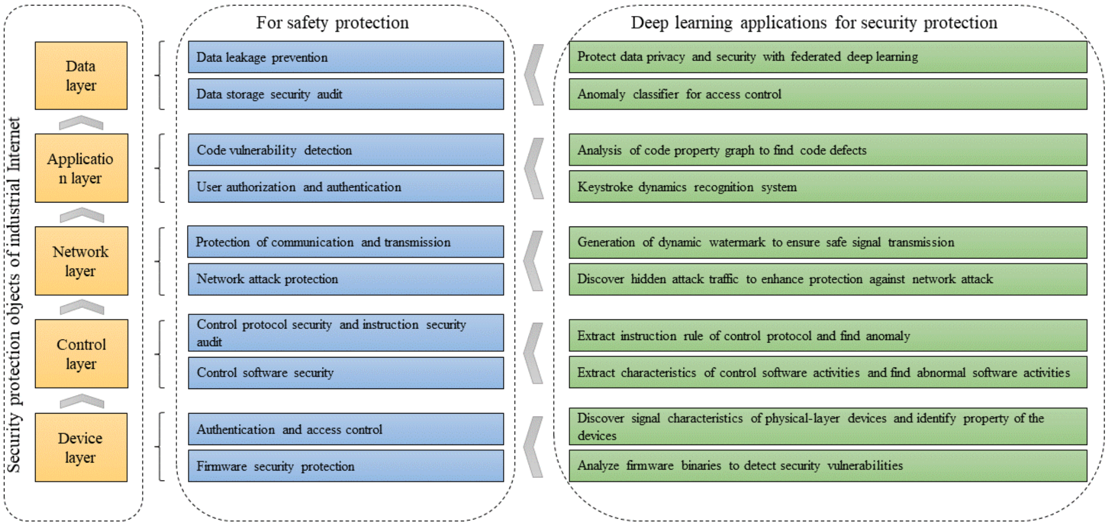
Security protection objects of industrial Internet
Data layer
Application layer
Network layer
Control layer
Device layer
For safety protection
Data leakage prevention
Data storage security audit
Code vulnerability detection
User authorization and authentication
Protection of communication and transmission
Network attack protection
Control protocol security and instruction security audit
Control software security
Authentication and access control
Firmware security protection
Deep learning applications for security protection
Protect data privacy and security with federated deep learning
Anomaly classifier for access control
Analysis of code property graph to find code defects
Keystroke dynamics recognition system
Generation of dynamic watermark to ensure safe signal transmission
Discover hidden attack traffic to enhance protection against network attack
Extract instruction rule of control protocol and find anomaly
Extract characteristics of control software activities and find abnormal software activities
Discover signal characteristics of physical-layer devices and identify property of the devices
Analyze firmware binaries to detect security vulnerabilities

Fig. 2. Classification of industrial Internet security based on deep learning.

3 Development status of deep learning applications for industrial Internet security

Industrial Internet security can be subdivided into security of various aspects, including device, control, network, application, data, and others [7]. The security requirements of each level and the application of deep learning are discussed below.

3.1 Deep learning applied to device layer security

Device security on the industrial Internet includes device identification, access control, and firmware security protection. The strong ability of deep learning to automatically and intelligently discover features and its powerful performance in binary analysis provides a new idea for identification and firmware code analysis of non-encrypted devices on the industrial Internet.

The openness of the industrial Internet means easy access to a large number of non-encrypted devices, leading to the vulnerability of corresponding devices to identity spoofing attacks. In other words, attackers can impersonate the identity of a legitimate device and send false information or take other malicious actions. Such attacks can be of great danger to key industrial devices and facilities and may cause physical damage. To prevent identity fraud, traditionally, encryption algorithms are used to verify the identity of the device. Unfortunately, many existing industrial Internet systems do not use cryptographic key operations. To illustrate, the Automatic Dependent Surveillance – Broadcast system, which is widely used in global aviation, does not have any cryptographic authentication, which means that significant investment will be necessary if cryptographic modifications are to be applied. Another feasible technique is to study the physical signals using an automatic encoder and a convolutional neural network. Devices connected to the industrial Internet can pick up subtle features at random during manufacture, and these features are reflected in the pulse-driven signals produced by the devices. Therefore, without knowing the specific characteristics of the signal device, the system can recognize the device, judge the property and identity of the device, identify all known devices, and report any unknown devices [8].

Because there are various industrial Internet platforms and firmware, firmware security is vital in the overall security architecture of the industrial Internet. To detect device firmware vulnerability, a common method is the binary similarity analysis of cross-platform firmware code, which aims to detect whether two binary functions from different platforms are similar; usually, this is performed with algorithms for approximate graph matching. However, this detection is slow, and misjudgment occurs when the differences between different instructions are few or subtle; thus, it would be difficult to apply this method to the industrial Internet, a field with high-level requirements for speed and security. In contrast, with a deep neural network, binary similarity analysis can be performed accurately and efficiently, since it can consider the graph embedding of binary code function segments as a neural network. It can compare the graph embedding proximity of two similar binary code functions, making its detection speed 3–4 orders of magnitude higher than that of traditional detection, thereby solving the problem of misjudgment presented by the traditional method [9]. Because of the above-mentioned advantages, deep learning technology can be used to detect binary code similarities and vulnerabilities, effectively supporting firmware security analysis.

3.2 Deep learning applied to control layer security

The control system of the industrial Internet connects upward to the network layer and downward to a multitude of industrial equipment, so protection measures to ensure its security are extremely important. Specifically, control security on the industrial Internet includes control protocol security mechanisms, command security audits, control software security reinforcement, and so on. As deep learning is capable of automatic feature discovery, it provides new insights into the detection of control protocol instruction attacks and control software.

The control system on the industrial Internet is divided into various subsystems of process control, data monitoring and acquisition, distributed control, and fieldbus control. These subsystems use control protocols to issue control instructions, so most attacks against such control protocols are performed by injecting incorrect data into control instructions in the process of transmitting the protocols. Traditionally, to detect attacks against commands, the common solution is to analyze the abnormal rules of attack messages and find similar attack behaviors. However, on the industrial Internet, where attack modes are constantly being updated, this method is not always reliable, as it may fail to detect new forms of attack. In contrast, the feature discovery capability of deep neural networks is expected to solve this problem: under normal control protocols, a deep neural network can learn the communication rules from the signals of the sensor and actuator acquired by process control devices, and then it performs security detection of the control protocol and instructions. In addition to detecting known instruction attacks, it can also identify new forms of attacks [10].

The control software of the industrial Internet faces security threats, such as malware injections. Common examples of such malware are fragments of elaborate computer programs intended to control and monitor industrial Internet assets that have been infected without detection. The traditional way to detect malware is to rely on manual labor to determine the characteristics of malware attacks. In other words, it uses characteristics that are already known to detect the software. However, this method is not feasible for polymorphic malware or virus detection. At present, many antivirus software suppliers have conducted in-depth research on employing deep learning methods to enhance the detection of malicious software, and these studies have achieved satisfactory results in testing in real-world situations [11].Therefore, it can be said that introducing deep learning technology into the control layer of the industrial Internet and using its inherent advantages of automatic feature extraction can provide the following benefits: dynamically analyze the characteristics of the industrial Internet control software activities, continuously analyze software activities and the execution of certain commands by software, detect behaviors of control software, and improve the ability of the control software to resist malware injection or other security threats [12].

3.3 Deep learning applied to network layer security

Security at the network layer of the industrial Internet includes communication and transmission protection, network attack protection, and other protections. At this layer, such abilities of deep learning as feature extraction, self-learning, and information compression can be utilized to provide new ideas for communication data encryption and detection of network intrusion on the industrial Internet.

On the industrial Internet, there are a large number of sensors, terminals, and control, computing, storage, and other devices. They need to transmit information about the surrounding environment and their own state, control instructions, and other information in real time and in a reliable and safe manner. This is especially the case in resource-constrained industrial Internet terminal nodes for which safe transmission of data can be a major challenge because its composition is relatively simple and its computing and storage capacity comparatively weak.Traditional transmission mode is highly reliable in that it depends on encryption algorithm, but at the same time, it also presents a relatively high level of complexity and delay when detecting attacks, so it is not suitable for large-scale deployment in an industrial Internet environment that requires low communication delay and complex composition. Therefore, a deep learning framework based on industrial Internet signals can be used, and random features (such as spectral flatness, skewness, kurtosis, and central moments) can be extracted from industrial Internet signals using long short-term memory (LSTM) to transform them into watermarks and load them in the original signal. Then cloud computing or edge computing nodes can be used to verify the watermark information to ensure the reliability of the signal, thus completing the network attack detection targeted at industrial Internet [13].

The industrial Internet is vulnerable to various targeted network attacks because of its complexity and sensitivity; therefore, it is necessary to configure an invasion detection system to scan network traffic activities and identify malicious or abnormal behaviors. Traditional intrusion detection systems usually use (shallow) machine learning technology, which cannot effectively solve the problem of intrusion classification and detection of massive amounts of data from environments that have real-time requirements. Deep learning is an ideal method for discovering hidden traffic and can be used to distinguish attacks from normal traffic. For example, the bidirectional long and short-term memory recurrent neural network method is used to explore in detail the network traffic characteristics of abnormal intrusions and to quickly and accurately identify abnormal activities, such as network attacks and network fraud against the industrial Internet [14].

3.4 Deep learning applied to application layer security

The application-level security of the industrial Internet includes user authorization authentication, code security, etc. The unique advantages of deep learning in understanding natural language and feature extraction provide new ideas for code security analysis and user authorization authentication.

The composition and functions of the industrial Internet are complex, involve much software, and have high requirements for the security of software source codes. Traditional code vulnerability detection primarily relies on manual analysis of code, knowledge of security issues, and experience accumulation of analysts, which makes it difficult to meet the analysis needs of the industrial Internet. A feasible way of thinking is to learn from natural language processing methods, take advantage of the unique advantages of deep learning in understanding natural language and memory function of natural language context of LSTM to comprehend and analyze the code attributes composed by source code abstract syntax tree, control/data flow chart, program dependency graphs, etc., as well as find and correct code defects in time during the source code programming stage, and actively complete code vulnerability analysis and detection [15].

The industrial Internet covers a wide area, requires high levels of security and privacy, and involves a large number of user authorization and authentication processes. Although traditional authentication systems based on passwords and personal identification codes are effective, they are not sufficient to resist many types of malicious attacks. Therefore, technologies such as face recognition developed by using the advantages of deep learning in biometric discovery have been successfully applied to application-level security [16], which plays a role in cooperating with traditional authentication systems and improving user authorization and authentication capabilities. In addition, to effectively improve the security of user authorization and authentication and reduce the cost, some researchers have proposed the use of deep learning technology on the keyboard to extract the time of each user keystroke, the pressure applied during typing, and the characteristics of the mobile device, touch area, and touch location to assist in user identification [17]. This solution provides a new approach for improving the authentication capabilities of industrial Internet users.

3.5 Deep learning applied to data layer security

One of the main tasks of industrial Internet data security is the prevention of data leakage. In the industrial Internet containing a large amount of fragmented data, reducing unnecessary cross-regional and cross-organizational raw data sharing and flow is important in improving data security; this is also an advantage of federated deep learning technology. In the federated deep learning system, self-owned data are not shared, and the parameters are exchanged through an encryption mechanism to establish a virtual shared model without violating data privacy protection regulations. Regarding data security auditing, industrial Internet data with low sensitivity can be stored in an industrial big data cloud platform at a relatively low cost. A data security audit mechanism based on deep learning is used to monitor data access and other behaviors to prevent data from being stolen, tampered with, and destroyed, thus ensuring data storage security.

The industrial Internet has multipoint communication needs connecting points, such as sensors, edge computing nodes, cloud end, and user end. In the traditional data processing process, the data are generated on the sensor side, and the preliminary collected data are stored in the relevant software [18]. Data intrusion and illegal access are mostly hidden in authorization and are not easily discovered. Unsupervised deep learning is used to classify abnormal behaviors to ensure that data are not stolen, tampered with, or destroyed. Perceived data are fed back to edge computing nodes, cleaned, preprocessed, and analyzed, then uploaded to the cloud end and further processed for users to extract. Sensitivity, fragmentation, and mass data flow are not conducive to data security protection in an industrial Internet environment. Therefore, to meet the multiparty computing requirements of industrial Internet data security, a feasible solution is to introduce federated deep learning technology to conduct data processing under the premise of not sharing sensitive data, minimizing data flow and unnecessary data transmission, and ensuring data security of the industrial Internet [19].

4 Challenges faced by deep learning applications

While deep learning technology envisages new prospects for industrial Internet security, vulnerabilities may be exploited by attackers and may be attacked by continuous advanced threats. For example, attackers can modify malicious files to circumvent deep-learning-based detection tools, add imperceptible noises to make the factory voice control system maliciously invoked, and paste small signs on traffic signs or other vehicles to mislead self-driving systems based on deep learning. In high-value or high-risk industrial production processes, if the deep learning system is maliciously attacked, it may cause equipment damage and even threaten personnel life or safety. There are five types of attacks against deep learning (Fig. 3): poisoning, model reverse, model extraction, physical, and adversarial attacks, which primarily occur in the model training and prediction stages.

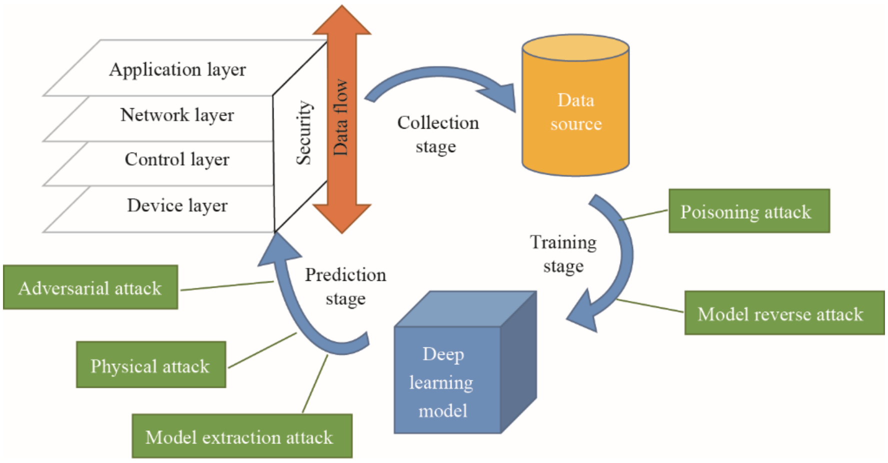

Fig. 3. Security challenges of deep learning system faced on industrial Internet.

DOI 10.15302/J-SSCAE-2021.02.013

4.1 Model training stage

A poisoning attack refers to attacking the training dataset to make the model unable to work properly. In the industrial Internet, competitors may manipulate training data by tampering with sensor measurements. For fault detectors based on deep learning, minor data tampering may lead to targeted misclassification or harmful behavior. A backdoor attack is also a poisoning attack. A special flag (backdoor trigger) is added to the training data process to bypass the classification of the model, such as adding a special code (text) to a file marked as non-malicious, and used for training deep learning models. The trained model can identify malicious files normally, but when a malicious file with this special code is detected, the model recognizes it as non-malicious, thereby avoiding detection. This attack does not affect the normal operation of the model most of the time and is extremely hidden [20].

A model reverse attack occurs during the completion stage of training. The information of the training dataset can be reversely extracted from the model through the output of the model (black box attack) and model parameters (white box attack); in other words, through the trained model data, the training data members of the model can be restored. Data are precious in the industrial Internet, especially sensitive data involving commercial value. For example, product parameters (such as weight, size, and model) required for the training of product quality detection models and sensor data in industrial production systems required by invasion detection systems involve product privacy and have certain commercial value [21].

4.2 Model prediction stage

In contrast to digital sample attacks, physical attacks are part of physical sample attacks, which deceive the deep learning model by changing the shape of the target object in real life or attaching a special mark. Physical attacks do not require tampering with the training data of the model, only passing a certain number of model function tests to discover the vulnerabilities and defects of the model to design and implement physical attacks. For example, the vision system of a self-driving car can use deep learning technology to classify pedestrians, vehicles, and road signs on the road, but after attaching a carefully designed note on the road sign, the vision system may not be able to correctly recognize the road sign. Moreover, glasses with special markings can interfere with proper working of a face recognition system based on deep learning. Under stable physical conditions, a targeted adjustment of posture, distance, and light can cause recognition errors in the face recognition system. There are many deep learning applications for face recognition and product quality inspection on the industrial Internet. If insiders are malicious, such physical attacks are relatively hidden, while threats are obvious [22].

Model extraction attacks refer to models with similar or even the same functions through public application programming interfaces (APIs). The specific parameters are difficult to grasp, and the purpose of the attack is to replicate models rather than restore data members [23]. It usually takes 20 to 30 days to train a model, and more complex models take longer. Some models applied to the industrial Internet have certain commercial value, such as abnormal behavior classification systems. These models have a certain degree of portability. If they are published on the Internet, even if only APIs are provided, criminals can obtain the input and output relationships through random combinations of inputs, thereby replicating models with the same function, making the company subject to loss of intellectual property rights.

Adversarial attacks are also called evasive attacks, which refer to the addition of some interference that is difficult for the human eye to detect in the normal sample, resulting in model prediction errors. They are divided into non-specific and specific target attacks. The user needs the model to judge a specific input as a specified output. At present, deep learning models are relatively vulnerable to adversarial attacks, and slight disturbances can interfere with the normal operation of the model. For example, the product detection model of the industrial Internet is vulnerable to adversarial attacks without a specific target. The attacker only needs to add noise that is invisible to the naked eye to the product image to make the model lose its detection capabilities. In severe cases, it can destroy the entire industrial production process. For a safety-oriented deep learning system, the attacker can bypass the security detection model by adding special sentences, thereby interfering with the safe operation of the industrial system [24].

5 Future research directions for applications of deep learning

5.1 The interpretability of deep neural networks

It is difficult for humans to understand the decision-making basis of deep neural networks because deep neural networks are usually used as “black box” models. Each neuron is obtained by superimposing a linear combination of the previous layer and a function, which is highly nonlinear. For industrial Internet security applications, people should also know the factors that the model is based on in addition to the final results of the model output. If the model is not interpretable, it means that the model itself is unknowable and insecure. Therefore, the prediction results of the model can be convincing only when the reliability of the information (such as not being attacked by poisoning, confrontational attacks, etc.) and the input of a clear model is ensured, the causality of input and output is identified, and deep learning can be applied to undertake the core tasks of the industrial Internet security system.

5.2 Sample collection and calculation cost

With the development of deep learning methods, the number of neural network layers is increasing in depth, and the number of training examples required and computing power requirements (power consumption) are also increasing rapidly. Even if the deep learning model achieves better results than traditional methods, the benefits of improving efficiency may not outweigh the increased costs. This will directly restrict the promotion and application of deep learning technology in industrial Internet security, which has diverse application scenarios requiring targeted collection of a considerable amount of data with manual labeling and high labor costs. As deep neural networks are large in scale, high performance is required to meet the requirements of accuracy and real-time function. The support of the computing system results in higher energy consumption requirements; therefore, it is necessary to study more efficient and automated dataset construction methods to lower the power consumption of deep learning models and computing systems.

5.3 Imbalance of the sample set

Deep learning has shown advantages in consumer Internet application; however, the application fields and scenarios of the industrial Internet are constantly changing, and it is difficult to provide a sufficient sample size for deep learning models. Therefore, it is necessary to study methods to increase the sample size using automated tools and deep learning methods based on small samples. In the face of fragmented, complex, and changeable industrial Internet security applications and scenarios, building a balanced sample set that can fully reflect the true distribution of data and using it to train deep learning models remains a challenge. At present, there are methods such as over- and under-sampling to alleviate the problem of sample imbalance in deep learning; however, practical systematic research results are still lacking [25].

5.4 Reliability of model results

The industrial Internet involves many fields, the frameworks of which are complicated, and the overall reliability of the system is very high. For example, the production of parts and components for aviation and aerospace aircraft requires equipment to achieve a reliability of not less than 99.999%. If important nodes fail, it will cause batch product damage or performance degradation. In the actual industrial production process, the stability of the model is more important than its expression ability. Once a problem occurs in a certain production chain, it may affect the operation of the entire production line. The prediction accuracy of many deep learning models is less than 90%, and thus it is almost impossible to transplant these models to industrial Internet applications that require extremely high reliability. Therefore, it is particularly important to study methods to improve the accuracy, certainty, and reliability of deep learning technology models.

5.5 Balancing availability and security

The application of deep learning has security and privacy issues. The training of deep learning models requires a large pool of data samples. While making a model public to realize its commercial value, the protection of the model and training data from illegal acquisition and use is a topic worthy of attention. The safety of the model is very important for industrial production. Some scholars have proposed ways to protect model privacy through differential privacy and homomorphic encryption methods and to detect adversarialities through adversarial training, thereby improving the security of the model. However, these methods reduce the usability of the model to a certain extent and affect its performance. Therefore, it is necessary to study the balance between the availability and security of deep learning models in the future.

6 Suggestions

6.1 Improve the overall security strategy of the industrial Internet

It is recommended to incorporate deep learning technology into the overall industrial Internet security strategy and build an in-depth industrial Internet security defense system that can cover the full lifecycle of security services and provide active and intelligent responses as the core feature. Deep learning should be reflected in the entire security architecture. Based on this, it connects all levels of the industrial Internet, establishes a dynamic defense system of "early warning, monitoring, handling, and protection" against security incidents, and maintains security systematically and comprehensively.

6.2 Overcome major problems in deep learning applications

In the application of deep learning in industrial Internet security, there are still some key technical issues that need to be solved urgently. It is recommended that researchers in computer science, neuroscience, automation, and other disciplines work together to develop application breakthroughs and aim at internationally leading development goals to build industrial Internet security ecology, proactively demonstrate the key directions of innovative research in cross-fields, and drive the deepening and expansion of the entire technology chain through demonstration effects. Based on the actual scenarios and urgent needs of industrial production, a joint-driven model of applications and problems should be adopted to tackle key technical bottlenecks in the integration of the two.

6.3 Guarantee the cross-integration investment of deep learning and industrial Internet security

Deep learning has broad application prospects and significant potential values in the field of industrial Internet security. Human capital, financial, and material investment in the cross-integration of deep learning and the industrial Internet should be increased reasonably. It is recommended to strengthen management policies or industry-specific planning research, encourage scientific research personnel to cooperate independently, deepen the tripartite cooperative relationship among industrial enterprises, universities, and scientific research institutions, form a political, industrial, academic, and research integrated cooperation system, and improve the deep learning technology system and its integration and application with industrial Internet security. Furthermore, the technology should be verified in practice to highlight its actual effects, thus forming a new pattern that promotes the development through scientific research.

References

[1] Li R Q, Wei S, Cheng Y H, et al. Research on typical application scenarios and standard system of artificial intelligence technology in intelligent manufacturing [J]. Strategic Study of CAE, 2018, 20(4): 112–117. Chinese.

[2] Li J H. Cyber security meets artificial intelligence: A survey [J]. Frontiers of Information Technology & Electronic Engineering, 2018, 19(12): 1462–1474.

[3] Amanullah M A, Habeeb R A A, Nasaruddin F H, et al. Deep learning and big data technologies for IoT security [J]. Computer Communications, 2020, 151(1): 495–517.

[4] Ha T, Dang T K, Le H, et al. Security and privacy issues in deep learning: A brief review [J]. SN Computer Science, 2020, 1(5): 1–15.

[5] Tsai S Y, Chang J Y. Parametric study and design of deep learning on leveling system for smart manufacturing [C]. Hsinchu: 2018 IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering (SMILE), 2018.

[6] Wang J J, Ma Y L, Zhang L B, et al. Deep learning for smart manufacturing: Methods and applications [J]. Journal of Manufacturing Systems, 2018, 48(C): 144–156.

[7] Yu X H, Liu M, Jiang X H, et al. Industrial Internet architecture 2.0 [J]. Computer Integrated Manufacturing Systems, 2019, 25(12): 2983–2996. Chinese.

[8] Liu Y X, Wang J, Li J Q, et al. Zero-bias deep learning for accurate identification of Internet of things (IoT) devices [J]. IEEE Internet of Things Journal, 2020, 11(4): 2627–2634.

[9] Xu X J, Liu C, Feng Q, et al. Neural network-based graph embedding for cross-platform binary code similarity detection [C]. Dallas: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017.

[10] Potluri S, Diedrich S. Deep learning based efficient anomaly detection for securing process control systems against injection attacks [C]. Vancouver: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), 2019.

[11] Kyadige A, Ethan M, Rudd B K, et al. Learning from context: A multi-view deep learning architecture for malware detection [C]. San Francisco: 2020 IEEE Security and Privacy Workshops (SPW), 2020.

[12] Kozik R. Distributing extreme learning machines with apache spark for NetFlow-based malware activity detection [J]. Pattern Recognition Letters, 2018, 101: 14–20.

[13] Ferdowsi A, Saad W. Deep Learning-based dynamic watermarking for secure signal authentication in the Internet of things [C]. Kansas City: 2018 IEEE International Conference on Communications (ICC), 2018.

[14] Roy B, Cheung H. A deep learning approach for intrusion detection in Internet of things using bi-directional long short-term memory recurrent neural network [C]. Sydney: 2018 28th International Telecommunication Networks and Applications Conference, 2018.

[15] Wang X M, Zhang T, Wu R P, et al. CPGVA: Code property graph based vulnerability analysis by deep learning [C]. Stockholm: 2018 10th International Conference on Advanced Infocomm Technology (ICAIT), 2018.

[16] Masi I, Wu Y, Hassner T, et al. Deep face recognition: A survey [C]. Parana: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018.

[17] Bernardi M, Cimitile M, Martinelli F, et al. Keystroke analysis for user identification using deep neural networks [C]. Budapest: 2019 International Joint Conference on Neural Networks (IJCNN), 2019.

[18] Yang C, Shen W M, Wang X B. The Internet of things in manufacturing: Key issues and potential applications [J]. IEEE Systems, Man, and Cybernetics Magazine, 2018, 4(1): 6–15.

[19] Yin B, Yin H, Wu Y L, et al. FDC: A secure federated deep learning mechanism for data collaborations in the Internet of things [J]. IEEE Internet of Things Journal, 2020, 7(7): 6348– 6359.

[20] Saha A, Subramany A, Pirsiavash H. Hidden trigger backdoor attacks [EB/OL]. (2020-07-15) [2020-12-15]. https://www.csee. umbc.edu/~hpirsiav/papers/hidden_aaai20.pdf.

[21] Hidano S, Murakami T, Katsumata S, et al. Exposing private user behaviors of collaborative filtering via model inversion techniques [J]. Proceedings on Privacy Enhancing Technologies, 2020 (3): 264–283.

[22] Boloor A, He X, Gill C, et al. Simple physical adversarial examples against end-to-end autonomous driving models [C]. Las Vegas: 2019 IEEE International Conference on Embedded Software and Systems (ICESS), 2019.

[23] Shafique M, Naseer M, Theocharides T, et al. Robust machine learning systems: Challenges current trends perspectives and the road ahead [J]. Design & Test IEEE, 2020, 37(2): 30–57.

[24] Wan M, Han M, Li L, et al. Effects of and defenses against adversarial attacks on a traffic light classification CNN [C]. New York: Proceedings of the 2020 ACM Southeast Conference, 2020.

[25] Buda M, Maki A, Mazurowski M A. A systematic study of the class imbalance problem in convolutional neural networks [J]. Neural Networks, 2017, 106: 249–259.