ISSN: 1813-4890

# Safety evaluation and verification of intelligent system

Jiayuan Li, Jiahui Pan, Changsheng Pan, Weikai Jing, Shiyou Luo

Software School, Northwestern Polytechnical University, Xi'an City, Shaanxi Province 710000, China.

## Abstract

**With the rapid development of big data and cloud computing, the application of artificial intelligence will become more and more popular. From smart and beautiful, convenient and portable intelligent equipment to payment methods that affect and change human life, driverless cars, and convenient transportation, the development of artificial intelligence has facilitated people 's life, work, and learning. At the same time, it has gradually brought many safety risks. This paper based on the analysis of the safety of artificial intelligence has 5 parts, the development history and researching status of artificial intelligence system; the future development trend of artificial intelligence system; the analysis for some potential safety hazards in the application and technology of the current artificial intelligence system; the corresponding countermeasures and suggestions for the existing security risks to establish a safety management framework, formulate safety guidelines and strengthen research on prevention technologies ; the research results.**

## Keywords

**Artificial Intelligence, Risks, Countermeasures.**

## 1. Introduction

### 1.1. Presentation of the question

At present, the increasingly complex social safety situation has been difficult to meet the requirements of substantive prevention and control of public safety. In recent years, the Chinese government has strongly encouraged the construction of Xue Liang Project. The integration of analog network signals and increased video surveillance has increased the scope of the monitor and strengthened and updated the protection and control system for public safety. High-definition cameras are widely used to compare pictures, face recognition, intelligent parking and other technologies and applications are popular to improve the safety of public services. It lays a foundation for improving the intelligence level of prevention, control system and social management. With the development of science and technology, more and more social fields have exerted artificial intelligence technology, which can provide more practical and personalized services for human beings. The technologies and applications not only bring us convenience, but also bring risks that cannot be ignored. Artificial intelligence system is not a complete system now [1].

### 1.2. International Researching Status

In the United States, Europe and Japan, artificial intelligence technology is developing very rapidly, leading the joint development of many related scientific fields of artificial intelligence systems. It is used in information science, control science, biological science, sociology, computer science and other technical fields. Artificial intelligence system is the trend of technical development. It is currently used in seven regions: personal assistant, security, self-driving, healthcare, e-commerce, finance and education [2]. We take the intelligent application domain video big data analysis in the security field as an example to analyze the big data in the

video. In a complex environment, we can realize a variety of information collection and event monitoring for people, vehicles and things, and have more accurate security guidance big data functions [3]. Application technology meets the needs of different scenarios and industries. At present, the development of artificial intelligence is an important strategy for all developed countries to protect the country and improve the overall strength of the country. In order to maintain the leading position in the international scientific and technological competition, each economy has its own development paths in the artificial intelligence system [4].

The UK is committed to developing artificial intelligence and industrial applications. Britain emphasizes that the effect of artificial intelligence is produced for development and comprehensive governance, which accelerates the transformation cycle of industry- university-research and enables the results to be implemented more quickly. The UK government also intends to spend about 200 million pounds of policy funds to establish higher education institutions for those who need high-skilled artificial intelligence training [5].

Institute. In addition to investment talents, artificial intelligence investment funds show an explosive growth trend. Research shows that investment in artificial intelligence increased significantly in 2009, attracting a total investment of $ 1.7 billion. Since 2013, more and more companies have joined the acquisition of artificial intelligence companies. In 2015 alone, 322 AI investment companies raised $ 2 billion. Statistics show that from 2011 to 2015, the level of funding for artificial intelligence increased by 62 % per year. In the future, growth will accelerate and not slow down. In general, the United States and Japan are ahead of other countries in artificial intelligence and investment technology [6].

## 1.3. Domestic Research Status

Since 1956, the development of artificial intelligence systems has had both lows and highs. In the theoretical side; there are three development directions: structural simulation, functional simulation and behavior simulation; from an academic point of view, there are three schools: artificial neural networks, expert systems and intelligent robots [7].

In July 2017, the State Council announced the development plan of a new generation of artificial intelligence, marking the beginning of the era of artificial intelligence in China. China has a large number of artificial intelligence talents in the eastern and central regions, and some cities in the west, such as Xi 'an and Chengdu, also have more related talents. The talents of artificial intelligence in the world are relatively concentrated in the fields of automatic learning, data mining and motivation identification, while the research fields in China are relatively scattered [8].

On July 10, 2018, Professor Li Kang proposed that "deceiving people" and "stealing people" are two major categories of artificial intelligence security issues. "Deception" is like taking a picture of a dog. Imagine that the data we got based on millions or even millions of data were directly used by others and used better than ourselves. This situation will cause incalculable losses to enterprises [9].

The report of the two national conferences in 2018 shows that the re- search and application of a new generation of artificial intelligence should be strengthened, which shows that China attaches importance to the development of artificial intelligence. At present, we need to solve some problems hidden in the development and dynamic analysis of artificial intelligence, and some related suggestions. So that artificial intelligence can develop healthily and effectively [10].

In 2017, the incident of face recognition cracking occurred in Shanghai Railway Station. Wang Qi took this as an example. By invading and controlling the attacker at the management end of the system, he could use the loopholes in the system to realize the access control of any face through the designated face recognition. At present, autonomous driving has become an important application field of artificial intelligence, but machine vision is not so reliable. Last

ISSN: 1813-4890

year, someone shared this case. In addition, the security of big data has also received much attention [11].

### 1.4. The purpose and significance of the study

Artificial intelligence is still an emerging technology under development. The defects of artificial intelligence technology will lead to the abnormal work of the system, which will lead to the emergence of security risks. For example, for example, the black-box mode will make the interpretability weak, and the development, design and production of intelligent systems will not start often, which will make the work abnormal. In addition, self-driving cars, robots and artificial intelligence devices without strict safety technical measures may be plundered and controlled, and we will be harmed if we accept the orders of criminals. Therefore, we should analyze the hidden risks of artificial intelligence system and put forward feasible countermeasures and suggestions.

## 2. The present situation and development trend of artificial intelligence system

At the beginning of 2016, when South Korean Go player Lee Se-dol lost his first game against Go program Alpha Go, artificial intelligence once again caught the public's attention and became one of the most striking topics of technology in that year. In recent years, even if the development of artificial intelligence is not as advanced as in movies, there is no denying that artificial intelligence technology has developed very rapidly, and it has gradually integrated into our life, work and study, and artificial intelligence is gradually growing [12].

### 2.1. The concept of artificial intelligence

Artificial intelligence is AI. The concept of artificial intelligence can be divided into two aspects: artificial and intelligent. Artificial intelligence means that artificial intelligence originates from human beings and transcends human civilization, and is the crystallization of wisdom of human development. Intelligence refers to the ability to simulate human behavior, thinking and thinking mode. The research of artificial intelligence computer science, its main purpose is to develop relevant theories and technologies in the near future, and to perform specific intelligent functions of human brain based on imitating machines.

### 2.2. Application of artificial intelligence in reality

Today's artificial intelligence system can help you complete complex mental work, logical thinking and other applications. Today, scientists have invented some electronic machines that can simulate human mental activities. The modernized and improved artificial intelligence system can help people complete difficult or even impossible tasks. The development of system or automatic robot has replaced some human activities, but it still can't reach the level of self-learning and upgrading in many aspects. There are also a large number of artificial intelligence products used for commercial purposes, such as customer information systems, decision support systems, medical consultants and legal consultants.

### 2.3. Development status of artificial intelligence

#### 2.3.1. Research and development status of intelligent interface technology

In order to achieve the purpose of man-machine communication, the intelligent interface technology is researched and developed. Therefore, scholars should combine theory with practice and make contributions to solving functional problems such as understanding and translating words and languages with computers and self-expression. At present, the rapid development of computer technology, as well as the great improvement in operating speed and

ISSN: 1813-4890

man-machine communication, are all due to the research and application of intelligent interface technology.

#### 2.3.2. Development status of data mining technology

The so-called data mining technology refers to the technology of classifying, mining and searching all kinds of fuzzy, unknown and potential data. Find useful data. Today, the three most important technologies, database, artificial intelligence and mathematical statistics, are divided into basic theory, discovery algorithm, visualization technology, knowledge representation method and semi-structure as research contents.

### 2.4. Development trend of artificial intelligence

#### 2.4.1. A driverless car

Many large automobile companies, such as Mercedes Benz and Toyota, are producing driverless cars. One of James Bond's films can independently understand the road conditions and drive by himself, which will soon become a reality. Artificial intelligence is not the only technology used in automatic cars. In addition, new technologies, such as automatic control and visual calculation, must be integrated to allow modifications to the existing vehicle structures and to automatically identify, analyze and control them.

#### 2.4.2. Intelligent classroom

Today, there are some intelligent educational software that allow students to provide textbooks, lectures and answer questions. Because students can communicate with teachers, intelligent classroom makes teachers' teaching both pleasant and comfortable. It is very convenient for students to check the mistakes in class at the end of the semester, and they can also see the learning materials they have studied in a few years.

## 3. Hidden Dangers and Defects of Artificial intelligence system

### 3.1. Analysis of potential safety hazards

The development of artificial intelligence is advancing by leaps and bounds. But in fact, intelligent image recognition technologies including visual recognition, face recognition and video scene analysis are far from safe and accurate application. In the application of artificial intelligence products, users generally lack self-protection awareness, and the safety prevention strategies of developers and builders are not fully deployed.

#### 3.1.1. Personal security

Artificial intelligence system still has decision-making bias that endangers the safety of human life. Artificial intelligence systems (especially highly autonomous systems) have specific decision-making and action capabilities. If cognitive deviation or network attack occurs, the system will make mistakes in judgment and take wrong measures, even undermining personal safety. For example, in November 2017, a self-driving bus that debuted on the first day of the Las Vegas highway collided with a truck two hours later. The bus is not technically responsible for the accident- according to the police, the truck driver complained that the bus was not sensitive enough in danger, just like the bus approached slowly, and he could not get rid of the danger

#### 3.1.2. Network information security

Mankind has experienced the era of network PC and mobile network, and is about to usher in the era of artificial intelligence. This era will focus on people and humanize all equipment. In the development of information technology, artificial intelligence is becoming more and more mature. In order to achieve the ideal goal, artificial intelligence needs to have a good information environment, which may pose a threat to information trans- mission applications when applicable [16].

ISSN: 1813-4890

### 3.2. Defects in artificial intelligence

Artificial intelligence is more accurate, more stable, more deeply rooted in people's hearts, has better vigilance and patience, and can work independently. From intelligent speakers, intelligent devices and intelligent robots to understanding translation, medical diagnosis and driving, artificial intelligence, which has become more and more common in our lives in recent years, has become more and more complicated. Nowadays, artificial intelligence is leading a new industrial revolution, which is closely related to each of us living and working. However, artificial intelligence is not perfect, which is not worrying. It has fatal defects and sometimes causes unexpected and serious consequences [17].

## 4. Countermeasures and Suggestions for Security Risks of Intelligence System

The three principles of robotics are the safety principles of robot design put forward by the famous science fiction writer Asimov. At that time, artificial intelligence had not yet appeared, but according to the core of the science novel Robot World, it adhered to this framework and became a loyal assistant and friend of human beings. Nowadays, the development of artificial intelligence brings great commercial potential and public worries.

### 4.1. Construction of artificial intelligence safety supervision framework

The recently published Report on Artificial Intelligence and National Security in the United States shows that artificial intelligence technology may become destructive technology in the future. In the national security fields such as nuclear weapons, airplanes, computers and life engineering, there is basically no Supervision Law and artificial intelligence institutions at present. There are few standards for evaluating the harm of artificial intelligence. In the United States, there are even some laws against unmanned cars and drones.

### 4.2. Develop artificial intelligence safety standards

The safety standard of artificial intelligence should be a collection of standards related to the safety, ethics and privacy protection of AI. Large- scale, with AI's own platform, technology, products and application security standards. Artificial intelligence is usually based on big data and usually faces various requirements for personal information protection, so the standards of big data security and personal information protection will play an important supporting role in AI security standards.

### 4.3. Strengthen the security research on artificial intelligence technology

The use of artificial intelligence in multidimensional vertical systems provides new opportunities for developing other industries. Human society wants to enter an era in which human beings and machines coexist. In order to ensure the controllability, solubility and coordination between artificial intelligence technology and human beings, various measures should be taken to solve the research and development risks. From technology, product design, mass production and popularization and application.

1. Improve the security of basic artificial intelligence technology, including sensors, network equipment and transmission security. The line confusion caused by artificial intelligence between the network world and the real world. In the fields of intelligent medicine and unmanned driving, it can cause physical loss and physical injury.

2. Improve the deep learning system of artificial intelligence. The concept of deep learning includes the study of artificial nervous system, the most representative learning method based on the data of mechanical learning. It forms a high-level abstract description by combining low-level features to express its attribute categories or characteristics.

ISSN: 1813-4890

### 4.4. Ethical design of artificial intelligence

In order to strengthen the safety of artificial intelligence application, the principle of ethical design must be observed in the design. Therefore, many scientists should pay attention to the ethical concept of artificial intelligence design involving human rights- related issues. The ethical concept design of artificial intelligence enables it to interact with human beings, establish the judgment ability and ethical ability of robots, and make the artificial intelligence system work within the defined framework. Applying moral concept to artificial intelligence system can improve the moral quality of robot and avoid the bad behavior of artificial intelligence system.

## 5. Summary

The rapid development of artificial intelligence will bring many un- known areas and new scenarios, but also bring unexpected security problems. It is necessary for the scientific and technological forces of all countries to join forces and jointly strengthen the research on the international common issues of artificial intelligence such as robot ethics and security risks. Strengthen international trade and cooperation, trade and cooperation with other aspects in laws, regulations, industrial image, etc., promote the unification of international artificial intelligence technical standards and safety standards, and provide more and better solutions for the healthy development of artificial intelligence to better serve people.

Artificial intelligence is a double-edged sword, but due to the increase of security threats, the development of human social history, advantages and disadvantages almost in all the technological revolution is co-existing, there will always be some debate, but the long-term debate, so that people can be well adapted to the technological innovation, gradually solve the potential risks, to achieve continuous improvement. Similarly, the development of artificial intelligence is historic and hailed as the core of the fourth Industrial Revolution. Global gross domestic product would grow by 14 per cent. In the end, security issues should be taken care of in the future, but they should not be a stumbling block to the development of AI.

## References

[1] Tom Simonite: When It Comes to Gorillas, Google Photos Remains Blind[EB/OL]. 2018.

[2] Xiumu.com The 7 major application fields of artificial intelligence will have earth-shaking changes in the future.2017-11-22

[3] LIN Jie Talking about the in-depth application of artificial intelligence in the field of security. China Public Safety.2017(6):116-119.

[4] Yi Shuihan. Comparative overview of the development status of artificial intelligence in China and the development status of artificial intelligence in the world. Heavy Qingshi Internet Industry Association.2018-5-19

[5] Ding Yubing Focus on foreign artificial intelligence development People's Daily.2017-06-08

[6] Analysis of foreign artificial intelligence industry development and application, which country is in the leading position Electronic Engineering World 2018-05-15.

[7] ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press.2016

[8] Chinese development status and future of engineering intelligence China Science and Technology Policy Research Center, Tsinghua University. Chinese Jing 2018-10-11

[9] Security is not paid attention to is the biggest hidden danger of artificial intelligence.IT Digital Home.2018-7-11.

[10] LIU Tingting, NIU Jinxing. Artificial Intelligence Development Trends and Security Challenges China Information and Communication Studies Yard 2018-10-19.

[11] Ren Xiaoyuan The safety of artificial intelligence has attracted attention Beijing Youth Daily.2018-8-3

ISSN: 1813-4890

[12] Zhang Daqu Development status and prospect of artificial intelligence. China Papers Network
[13] Liu Tingting, Niu Jinxing Artificial Intelligence Development Trends and Security Challenges. China Information and Communication Studies Yuan.2018.10.19
[14] Lü Zeyu The history, present and future of artificial intelligence Information and Computer.2016:166-167
[15] Lu Yan Safety risks of artificial intelligence and future employment. China New System Department Research Institute.2018-0-18
[16] LIU Shuchang . Artificial Intelligence Security Problem and Its Solution [J/OL]. Electronic Technologies. 2018(15):246
[17] Regarding the flaws of artificial intelligence that cannot be ignored Gale.2018-08-29