Application of Artificial Intelligence Voice Technology in Radio and Television Media

Meihui Hu, Zhiwei Xiang, Kai Li*
State Grid Xinjiang Electric Power Co., Ltd. Information and Communications Co

*Corresponding author’s e-mail: 1158852782@qq.com

Abstract: With the application of artificial intelligence in various fields, more and more people see the contribution of artificial intelligence technology to the development of the industry, and put more energy in the research of artificial intelligence language, and strive to provide better technical support for the development of more industries. In radio and television media, artificial intelligence voice technology can play a very important value, it can effectively improve the efficiency and quality of traditional audio work, optimize the singing system, broadcasting system and retrieval system, so as to provide better service quality for the masses. This paper discusses the application and development of artificial intelligence technology based on the fusion media environment, analyzes the application and development of artificial intelligence technology in promoting production efficiency, intelligent robot writing, intelligent face recognition, intelligent speech semantic recognition, intelligent OCR recognition, automatic broadcast and other aspects. These AI applications provide efficient and secure support for services, greatly improving service efficiency.

1. Introduction
In recent years, artificial intelligence technology has developed rapidly, among which intelligent voice technology is the most outstanding. At present, it has realized the functions of fast information input and convenient man-machine interaction, which can convert speech recognition into text, convert text into voice through speech synthesis, understand the meaning of speech and execute corresponding commands and operations. At present, ai voice technology has been widely used in government and enterprise, education, medical care, automobile, finance and other fields [1]. In the future, the development of radio and television should make use of scientific and technological innovation to promote the intelligent development of radio and television program production, content production, safety supervision and other aspects. The rapid development of artificial intelligence in recent years has provided possibilities for the future development of radio and television [2]. As a brand new technology, intelligent voice technology can realize functions such as rapid input of information and human interaction, and is currently involved in many industries such as education, medical care and automobile[3]. Although the development of science and technology has put forward new development goals, and in the process of the development of radio and television media, a large number of documents and audio need to be processed, but the use of intelligent artificial voice technology in practical work is not common[4]. The traditional manual processing not only takes a long time, but also works inefficiently. If the manual method is careless, it will cause the deviation of the content understanding and the problem of recording. How to make good use of the program material library of the radio station, make it play a greater role in the radio program production, audio resource sharing and public service, is the urgent demand of the radio station business development[5]. One of the difficulties of effective
reuse of broadcast program material library lies in providing powerful and efficient retrieval function. In the traditional program material library management method based on metadata cataloging, audio material is different from text resources, and its content is difficult to express in words[6]. Simple primary catalog information to the description of the material is too rough, cannot achieve the requirements of the detailed content retrieval, senior catalog and need a lot of manpower, radio in the face of huge amounts of audio input on library is too big, even if done shows, clips, scenarios, such as catalog, also can't fully meet the diverse and personalized retrieval requirements for content[7]. This requires true automated processing based on content understanding and artificial intelligence to automatically index and process audio to provide more detailed and personalized content search. In this paper, the actual application of artificial intelligence voice technology in radio and television media is discussed, and a new intelligent voice application platform for radio and television media is built on the basis of traditional manual processing and learning from experience[8].

2. **Intelligent voice technology classification**

2.1. *Speech recognition technology*

Speech recognition technology mainly refers to the spoken speech into the corresponding text for input, is a new type of information input technology. Speech recognition technology first to detect the endpoint of the audio, to determine the position of speech, and then recognition processing. Intelligently judge the punctuation and other relevant information of the input information by recognizing the dialogue context of the statement. In the process of input, the system selects the corresponding words preferentially by identifying the user-defined words[9].

![Speech recognition pattern diagram.](image)

Table 1. Speech transcribe function

| Type               | Function                                                                 |
|--------------------|--------------------------------------------------------------------------|
| Phonetic transcripts | The reporter forms the program manuscript through the online multi-terminal input voice |
| Intelligent lyrics  | Program dubbing or concurrent audio automatically generated with time code of the word file |
| Dialogue to text    | Translate the conversation content of guest host in studio interview program into text in real time |
| Clues to the sorting| The audio files gathered by financial media are roughly sorted by keywords |
| Audio retrieval     | The reporter forms the script by inputting the voice in multiple terminals |
| Material catalog    | Material files through the intelligent voice transliteration system to generate "text business cards" as the cataloging basis |
2.3. Speech synthesis technology
Speech synthesis refers to the output text, the text is processed by the computer, captured into fluent text and sentences, but also can use this stage of the language to play, language navigation and voice prompt many applications. The main content of language combination is to judge the order of words, to data-process words, to derive language wave signals, and finally to put the unit sequence into waveform series, and finally to output the complete signal[11].

3. Technical architecture of artificial intelligence speech recognition technology

3.1. System architecture
From the perspective of the composition of the system architecture of ai speech recognition platform, it is composed of five parts: basic support layer, core capability layer, application service layer, application access layer and service application platform[12-17].

| Architecture                | Function                                                                                                                                                                                                 |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Foundation support layer    | The basic support layer mainly uses cloud computing to manage the hardware and storage resources at the bottom of the platform.                                                                           |
| Core competency layer       | Be able to utilize the core engine for practical applications of radio and television media                                                                                                |
| Application services layer  | The application service layer can integrate the latest voice technology to provide powerful voice services and other related functions for platform management and service.                                      |
| Application access layer    | The application access layer can realize service interface and voice processing, and can provide intelligent voice services.                                                                                  |
| Service application platform| The service application platform can realize the related functions such as script singing, virtual broadcasting, content transliteration and so on, and support the docking work of third-party services. |

3.2. Software architecture
There are five main software architectures in ai speech recognition technology. The first is the Java application specification architecture composed of J2EE. Such systems are portable across systems, allowing for highly flexible and scalable system design by deploying them on related application servers. The second is to use nod-webkit as an application program in the client, providing an interactive interface to facilitate the use of business personnel. Administrators can view the usage from any place, so that the security of data can be guaranteed reliably. The third enables restful interfaces to improve response speed, performance, efficiency, and ease of use. The fourth is the data exchange format designed by JSON, which is a language text format that is easy to read and write. It can be used across platforms, across systems and across languages, with strong versatility, flexibility and expansibility. The fifth is the Spring Boot framework, a framework for easy deployment and configuration that provides XML configuration and quasi-production application monitoring[18].

4. Exploration of the application of artificial intelligence in broadcasting field

4.1. content level
As ai plays an important role in modern society, it is also popular in the broadcasting field. In the era of big data, effective use of broadcast content is conducive to the development and mining of big data. At the same time, it also provides convenience for the broadcasting field, and people can acquire information quickly in the era of big data. As a result, radio transmission was quickly accepted by the
public. For example, traffic information plays an important role in broadcasting. At the same time, people attach great importance to the content of traffic, and people with different jobs will have different opinions on the access to traffic information. While driving, most drivers need to focus on personal safety and access traffic information, so they choose the more traditional way. At this time, artificial intelligence will come in handy. Artificial intelligence can identify some traffic problems through traffic cameras, so as to feedback these problems to the people, so that the public drivers can effectively avoid these traffic accidents, so as to protect their own safety. Or artificial intelligence can effectively broadcast the content of traffic through the way of broadcast communication, so that big data mining can have an in-depth understanding of the broadcast communication content. In order to protect the popularity of AI in the radio station, the radio staff and designers can do some image design for AI, so as to improve the image of AI in people's mind[19].

4.2. Form level
In order to ensure the sustainable development of new media, enterprises need staff and managers to have a deep understanding and analysis of the interaction between users and platforms. Because interaction is an essential part of the user experience, the field of broadcast communication has evolved because of interaction. Therefore, the use of AI in broadcasting stations can effectively improve the efficiency and speed of broadcasting, and staff and technicians can receive regular training to improve the rational use of AI technology. The management personnel of the radio station can discuss and analyze the meeting, and create a virtual host for the radio station by using artificial intelligence technology, which effectively increases people's attention to the radio station and provides good basic conditions for the development of the radio station. For example, designers and technicians can use artificial intelligence technology to make the host and virtual host reasonable interaction, add to the wonderful program. Thus, the sustainable development of the radio station is guaranteed, and the project is set up so that customers can communicate with the virtual host by voice, which improves the enthusiasm of the masses and improves the formalization of the radio station[20].

5. Application of artificial intelligence speech recognition technology in Radio and Television media

5.1. Script singing system
With the script singing system, editors can quickly convert the content dictated by audio files into scripts. It should be noted here that the non-Mandarin text requires the reporter to use the microphone to read after the recognition. For some dialects, the system can use the dialect acoustic model package for identification. On the one hand, it improves the efficiency of text recognition, and on the other hand, it solves the problems of recording positioning and correction. The system can reduce the noise of the audio files recorded in the party and solve the problem of clause and subsection[21].

5.2. Content transliteration system
Content transfer system is mainly to not enter the audio content of medium resource library As language content for media files to generate their own "word card", then the media file with the card, submit audit so cataloging personnel can through "word card" for more information about the audio file, cataloguing work better, Reduce the error rate of cataloging work[22].

5.3. Virtual broadcast system
Virtual host, intelligent voice robot and the main broadcast library together, in the use of the process of the document will be imported can be dubbing work, realize the automatic broadcast news function, in the broadcast process, can achieve the speed, loudness, voice template control, in order to adapt to the requirements of specific occasions. The virtual host of the second element mainly uses the virtual host of cartoon to broadcast automatically, which innovates the broadcast form on the basis of the underlying language synthesis service. The speech interaction robot can dictate in Chinese, and the recognition
rate of man-machine dialogue is more than 90%, which can provide dialogue content for the needs of radio and television[23].

5.4. Third-party interface services
The platform provides the third party with synthesis, speech recognition and other business functions by providing the development interface in the application. Other systems can connect data resources to the voice cloud platform and use intelligent financial media to realize the sorting function of clues. Through the way of convergence of financial media clues, the voice can be formed into text documents, and these text contents can be classified according to the needs of users, so as to provide users with subscription clue services. It is widely used in topics planning, production and release of content[24].

5.5. Integrated media conference system
The financial media conference system is mainly used in the meeting minutes of radio and television media. In the meeting scene, the collected information is converted into text in real time, and the user can sort out the meeting key points according to the need. In addition, you can import a recorded audio file and convert it into text. During the transcription, the key points of the meeting are automatically marked. You can double-click the text to listen back to the related audio file, so that the audio is corresponding to the text. By filtering the modal words and segmenting the content, the recorder can understand the audio file more effectively. Users can also export the meeting template to form the meeting minutes through the functions such as search and query[25].

Table 3. Application and advantages of AI speech recognition technology in radio and television media

| Application                        | Advantages                                      |
|-----------------------------------|-------------------------------------------------|
| Script singing system             | Improving the efficiency of text recognition and solving the problems of recording positioning and correction. |
| Content transliteration system     | Simplifying work intensity of cataloging personnel and reducing the error rate. |
| Virtual broadcast system          | It can adapt to the requirements of specific occasions with high recognition rate. |
| Third-party interface services    | Providing users with subscription clue service with intelligent classification. |
| Integrated media conference system| Convenient and fast, real-time meetings and other into text. |

6. Application of artificial intelligence in digital media
Today, with the rapid development of digital media technology, artificial intelligence technology has been widely used. In digital media, artificial intelligence has a very wide range of applications, it can constantly optimize the role and efficacy of digital media, but also can improve the overall advantages of digital media application, to meet people's needs in an all-round way.

For example, as a traditional media, sci-tech periodicals have begun to attach importance to cross-media integration and try to improve the communication efficiency of sci-tech periodicals by means of new media.

As shown in Table 4, there were reading status of four wechat public accounts opened in 2014. The frequency of news push of architectural Journal journal of Tongji University was significantly higher than that of Textile Journal and Journal of Mechanics. According to the total reading volume and average message reading data, the wechat official accounts of Architectural Journal and Journal of Tongji University both have high reading volume. With the continuous improvement of artificial intelligence technology, the value of the public account has been highlighted. The customized menu function of
wechat official account based on artificial intelligence technology can realize the integration of wechat and official website, wechat customer service, wechat and publishing, and other functions. Artificial intelligence technology enables wechat public account to achieve the integration of mobile terminal, Internet, print media and other media.

Table 4. reading status of four wechat public accounts opened in 2014

| Public account                  | Reading Message number | Average reading | Push interval |
|--------------------------------|------------------------|-----------------|---------------|
| Textile journals               | 2149                   | 27              | 79.6          | 13.5         |
| Journal of mechanics           | 577                    | 9               | 64.1          | 60.8         |
| Journal of architecture        | 85036                  | 42              | 2024.6        | 8.7          |
| Journal of Tongji University   | 201731                 | 110             | 1833.9        | 6.7          |

As can be seen from Table 5, the most popular content for wechat public account fans is experience and sharing, followed by content guidance. Among them, Architectural Journal takes the public account as the new media communication platform of its paper journal, and uses artificial intelligence technology to push the content mainly focuses on the wonderful contents of the journal. The content pushed is exquisite, full of pictures and pictures, and is welcomed by fans.

Table 5. The number of different types of messages pushed and read

| Public account                  | Content introduction Message number | Content introduction Reading | Industry dynamics Message number | Industry dynamics Reading |
|--------------------------------|-------------------------------------|-----------------------------|---------------------------------|---------------------------|
| Textile journals               | 2                                   | 7                           | 7                               | 566                       |
| Journal of mechanics           | 7                                   | 494                         | 0                               | 0                         |
| Journal of architecture        | 35                                  | 70372                       | 6                               | 11512                     |
| Journal of Tongji University   | 4                                   | 21                          | 0                               | 0                         |

| Public account                  | Experience Sharing Message number | Experience Sharing Reading | Notices and Announcements Message number | Notices and Announcements Reading |
|--------------------------------|-----------------------------------|----------------------------|------------------------------------------|-----------------------------------|
| Textile journals               | 1                                 | 3                          | 0                                        | 0                                  |
| Journal of mechanics           | 0                                 | 0                          | 1                                        | 27                                 |
| Journal of architecture        | 0                                 | 0                          | 0                                        | 0                                  |
| Journal of Tongji University   | 10                                | 111                         | 87                                       | 20192                              |

7. A database of ai voice technology

7.1. Collect information
The daily work of radio and television media includes program production, contentcataloging, program broadcasting and meeting minutes, etc. In order to improve the transcription effect of these work, it is necessary to collect relevant useful information and optimize the speech transcription in a scientific and
targeted way to meet the requirements of intelligence. The corresponding information resource database can be formed on such a base, which can be stored effectively[26].

7.2. Requirements on the database specification
Database in the use of related requirements specification, mainly including the following aspects: first, can support multiple users in online business processing, full of multiple users to access the relevant requirements of the database. Second, the use of technology has high reliability, and in order to ensure better results, it is necessary to have high fault tolerance and self-recovery capability. Third, the use of distributed and heterogeneous data sources has corresponding support, data can be automatically replicated and data synchronization between multiple copies should be done well. Fourth, the corresponding security management mechanism can be provided. Fifth, it can provide a set of software development tools that support the entire cycle of software development use[27].

8. Conclusion
In view of the practical application of artificial intelligence voice technology in radio and television media at the present stage, the practical application of this new technology in media is not much, and needs further research and discussion. In general, the use of artificial intelligence voice technology in radio and television media is mainly in the aspects of script singing, virtual broadcasting, content transcoding, media meeting and third-party interface services. The use of these functions greatly improves the working efficiency of radio and television media and the traditional working efficiency. In the future development of radio and television media, more attention should be paid to the use of artificial intelligence voice technology, and it should be reasonably used according to corresponding norms. Only when it is properly used can it promote the further development of radio and television media, keep up with the pace of new media development, and promote the whole industry to a higher level of development.

Reference
[1] Kumar M, Sharma S, Chaudhary D, et al. Image Recognition Using Artificial Intelligence[C]// 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). 2021.
[2] Chen Y, An Q. Brief introduction of face image recognition method based on artificial intelligence[C]// Third International Conference on Image, Video Processing and Artificial Intelligence. 2020.
[3] Wang M. Bronze Culture Image Recognition System based on Artificial Intelligence and Network Technology[J]. Journal of Physics Conference Series, 2020, 1574:012097.
[4] Chen R, Wang M, Y Lai. Analysis of the role and robustness of artificial intelligence in commodity image recognition under deep learning neural network[J]. PLOS ONE, 2020, 15(7):e0235783.
[5] J Zhang, Li S L, Zhou X L. Application and analysis of image recognition technology based on Artificial Intelligence -- machine learning algorithm as an example[C]// 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL). 2020.
[6] Lu J, Xiong S, Wang M. Application of Image Recognition Technology Based on Artificial Intelligence in Traffic Meteorological Service[J]. China Computer & Communication, 2019.
[7] Wang K, Wang X, Na M A, et al. Application Research of Image Recognition Technology Based on Artificial Intelligence in Pumped Storage Power Station[J]. Hydropower and Pumped Storage, 2019.
[8] Hu R, Li M, Xu H, et al. Research and Application of Key Technologies for Medical Image Intelligence Knowledge Discovery and Data Processing[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2020.
[9] Zhong X, Shih F Y, Guo X. Automatic Image Pixel Clustering based on Mussels Wandering Optimization[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2020.
[10] Yang X, Wang X, Ma Y. Simulation of Polarization Image Based on Intensity Image[J].
International Journal of Pattern Recognition and Artificial Intelligence, 2020.

[11] Duan R, Wang Y, Qin H. Artificial intelligence speech recognition model for correcting spoken English teaching[J]. Journal of Intelligent and Fuzzy Systems, 2020, 40(1):1-12.

[12] Lee J H, Huh T R, Han H W. Discourse Structure of Artificial Intelligence Speech Recognition Service[C]// Advanced Science and Technology 2018. 2018.

[13] Kularski C M. Artificial Intelligence Speech Recognition with Paralinguistic Features and Absent Context. 2015.

[14] Nilsson N J. The Quest for Artificial Intelligence: Speech Recognition and Understanding Systems. 2009.

[15] Graupe, Daniel. Artificial intelligence pattern-recognition-based noise reduction system for speech processing[J]. Journal of the Acoustical Society of America, 1993, 93(5):3018.

[16] Ghaidan K A, Issa H, Trad E. Artificial Intelligence for Speech Recognition Based on Neural Networks[J]. Journal of Signal & Information Processing, 2015, 06(2):66-72.