

# ***Research on the Impact of ASR Technology on Intelligent Interconnection Based on Embedded Experiments***

**Xinwei Liu<sup>1,a,\*</sup>**

<sup>1</sup>*Art and Design college, Shenyang Ligong University, Shenyang, Liaoning, China*  
*a. 1319liu2678@asu.edu.pl*

*\*corresponding author*

**Abstract:** With the rapid development of artificial intelligence technology, automatic speech recognition (ASR) technology has become an important breakthrough in the field of human-computer interaction. ASR technology can convert human speech into text, thereby revolutionizing the way human-computer interaction occurs and significantly improving the efficiency of information processing. Therefore, it is of great research significance to observe the impact of ASR technology on social interaction and public discourse space from the perspective of embedded experiments. This article comprehensively ASR technology, covering its working principle, development history, technical methods and application scope. The basic principles, application scope, and related technologies of embedded experiments are introduced, and the direction of future development is prospected. In addition, this article explores the specific application of embedded experiments in ASR technology, analyzes in detail how ASR technology is combined with embedded experiments through steps such as preprocessing, feature extraction, and identification, and gives three specific cases to illustrate its use. Practical application effect. The research significance of this article is to reveal the important role of ASR technology in various fields, and how embedded experiments can further promote the optimization and progress of ASR technology.

**Keywords:** ASR, Embedded Experiments, Human-computer Interaction.

## **1. Introduction**

With the rapid development of artificial intelligence technology, automatic speech recognition (ASR) technology has become an important breakthrough in human-computer interaction. ASR technology converts human speech into text, promoting the innovation of human-computer interaction and greatly improving the efficiency of information processing [1]. From the perspective of embedded experiments, observing the interaction between ASR technology and society and evaluating its impact on the public discourse space is a research topic with great practical significance.

ASR is an interdisciplinary field that integrates disciplines such as linguistics, psychology, signal processing, acoustics, pattern recognition, artificial intelligence and machine learning [2]. Through preprocessing, feature extraction, acoustic modeling, language modeling, recognition and post-processing, it converts human speech into computer-readable language, breaking the barrier between humans and computers. In recent years, ASR technology has been widely used in many fields such as the Internet of Things, smart home, medical health, and industrial fields. In the Internet of Things, ASR technology enables devices to interact through voice commands, improving the intelligence

level of the system; in the smart home field, people can control home appliances through voice and realize intelligent control of the home environment; in the medical and health field, ASR technology assists doctors in recording medical records and diagnostic information, improving work efficiency; in the industrial field, ASR technology promotes the development of intelligent manufacturing and remote monitoring. Through embedded experiments, researchers can delve into how ASR technology changes how people participate in public discussions. ASR technology has shown great potential in the Internet of Things, smart homes, medical and other fields. By improving the processing efficiency and accuracy of voice information, it has significantly improved the intelligence level and user experience of devices in various fields, breaking language and other boundaries. Barriers to achieving seamless communication between different languages and humans and computers.

This article systematically analyzes the theoretical basis, development history, related technologies and applications of automatic speech recognition and embedded experiments in multiple fields, and deeply explores the implementation method of combining automatic speech recognition and embedded experiments.

First of all, the article starts with the basic theory and provides a comprehensive introduction to ASR, including its operating principles, work processes, development stages, technical methods, etc. In addition, the article also discusses the current application status of ASR in different fields, such as smart homes, voice assistants, vehicle systems, etc., demonstrating the broad application prospects of ASR technology.

Then, in the embedded experiment part, the article first introduces the basic theory and architectural design principles of embedded systems, emphasizing the particularity of its operation in resource-constrained environments, as well as the development environment and hardware platform. Secondly, the application scope and updates of embedded experiments are introduced. Finally, the article introduces the application cases of ASR combined with embedded systems. Through the analysis of multiple practical application cases, the article shows how to use embedded ASR technology in actual scenarios to solve specific problems and improve the intelligence level of the system and user experience. At the same time, the article also points out the current challenges and future development directions, providing reference and inspiration for subsequent research and application.

Through systematic analysis and detailed elaboration, this article aims to provide readers with a panoramic view of the application of ASR combined with embedded systems, which is of great significance for promoting the development of human-computer interaction technology and improving the level of social informatization.

## 2. ASR Basic Information

ASR is one of the ways to convert acoustic signals into text [3]. Its main working principle is: that the sound acquisition device captures the signal, and the preprocessing module performs noise reduction, filtering, digitization and other processing on the collected sound signal, the feature extraction module converts the processed sound into recognizable features (such as spectrogram, phonemes, etc.), and then the recognition module compares the extracted features with the pre-stored model, converts the voice signal into text, and finally performs grammar checking and error correction, and finally outputs the result to realize the voice interaction function.

Automatic speech recognition is a key branch in the field of contemporary artificial intelligence. Its development momentum is rapid and it is becoming a hot technology. Automatic speech recognition technology focuses on converting human language into text or instructions, so as to achieve deep interaction with computers and smart devices.

In the process of development, ASR has gone through several stages. Early ASR technology mainly faced problems such as low recognition accuracy, insufficient efficiency and limited scene

application. Due to the immaturity of the technology, the early system performed poorly when dealing with complex accents and background noise, resulting in a low recognition rate. In addition, the application scope of these technologies is limited to specific scenarios, such as voice dialing and simple voice commands, which makes it difficult to meet diverse needs. Nowadays, with the continuous advancement of technology, the performance of ASR systems has been significantly improved. They can accurately recognize speech content with different accents and noise, while privacy security and processing efficiency have also been greatly improved. For example, by replacing cloud servers with edge computing devices, SAR systems can achieve localized processing and effectively protect user privacy [4, 5]. In addition, the application of ASR systems has expanded from a single scenario to a variety of fields, providing a convenient and effective voice interaction experience for various fields.

ASR technology is mainly divided into traditional methods and deep learning methods. Traditional methods are based on pattern matching and statistical models, and usually use Fourier transform, hidden Markov models (HMM), etc. to extract features; deep learning rules are based on neural networks and integrate context by training multi-layer recurrent neural networks (RNNs) to improve accuracy. At present, deep learning methods have become a mainstream trend. They are far better than traditional methods in terms of accuracy and convenience. At the same time, they can effectively recognize speech emotions and have broad application prospects [6, 7, 8].

With the continuous advancement of technologies such as neural networks (CNN), the accuracy of automatic speech recognition systems has been significantly improved, and the application fields have also continued to expand, involving smart homes, smart vehicles, mobile communications, medical care, finance, etc. fields. In the field of smart home appliances, voice recognition is used to control the switches of home appliances. Install voice intelligent navigation, voice control vehicles, etc. in the car. Speech recognition cases, voice control of medical equipment, voice-assisted diagnosis and monitoring, etc. in the medical field. In the field of mobile communications, ASR technology can realize speech recognition input, speech recognition translation and other functions. In some service industries, customers' voices can be recognized and converted into text to improve work efficiency. In the future, with the continuous advancement of artificial intelligence technology and the expansion of application scenarios, ASR technology will play an important role in more fields. Future trends may include further improvements in recognition accuracy, the popularization of multilingual support, and other artificial intelligence-integrated applications of technology, etc. [9].

ASR technology is an important branch in the field of contemporary artificial intelligence, with rapid development momentum and broad application prospects. With the continuous advancement of technology and the continuous expansion of application scenarios, ASR technology will play an important role in more fields and bring more convenience to people's lives and work.

### 3. Embedded Experiments

Embedded experiment is a comprehensive research method that integrates specific modules into the actual operating environment for in-depth testing and analysis. This experimental method evaluates the effectiveness and performance of different technologies or strategies by comparing their performance in diverse user groups. Embedded experiments are usually based on a certain foundation in user data mining and analysis, to provide strong support for technological innovation, user experience optimization and system performance improvement through detailed data collection and processing analysis [10]. At the same time, embedded experiments rely on specific development environments and hardware platforms, such as Raspberry Pi, Arduino, etc., as well as various microcontrollers (MCU), digital signal processors (DSP), graphics processing units (GPU), application-specific integrated circuits (ASIC) and field-programmable gate arrays (FPGA).

The basic principles of embedded experiments mainly include system integration, environmental simulation, data collection, and data analysis. For example, in a smart home system, embedded experiments may involve integrating sensors, controllers, and communication modules to simulate various usage scenarios in a home environment, collect data, and analyze the performance of the system under different conditions.

With the development of technology, embedded systems are widely used in various fields to improve performance through interconnection, data exchange, and intelligent control [11]. Embedded experiments can be used to evaluate the effects of technologies or strategies in consumer electronic products such as smart homes and smart wearable devices, such as speech recognition and human-computer interaction. In the field of medical devices, they can be used for the development and optimization of medical devices, such as real-time monitoring, remote monitoring, and intelligent diagnosis and treatment [12]. In industrial automation and transportation systems, embedded experiments also play an important role in improving productivity and safety by optimizing the operating efficiency of production lines and equipment.

The technology involved in embedded experiments is also constantly updated and upgraded. For example, embedded experiments will pursue more efficient processors such as Raspberry Pi 4 and the latest ARM processors, which provide more powerful computing capabilities. Miniaturized hardware, such as microcontrollers and single-board computers, make embedded systems more compact and energy-efficient. Advanced data analysis tools use machine learning and artificial intelligence technology to improve data processing and analysis capabilities. Edge computing processes data locally on the device, reducing latency and bandwidth requirements and improving system response speed [4].

In the future, embedded experiments may deeply integrate artificial intelligence and machine learning technologies to significantly improve the level of intelligence and achieve more efficient and accurate decision-making and automated control. At the same time, system performance will be greatly improved to meet increasingly complex computing needs through the use of multi-core processors and increased storage capacity. In terms of power consumption, low-power design and energy-saving technology will become mainstream to extend device battery life and improve user experience. These technological advances will further promote the development of embedded experiments and make their applications more extensive and in-depth in various fields. In general, the embedded experiment is a comprehensive research method to evaluate the effect and performance of technology or strategy. Its application involves multiple fields, which can help to analyze the user experience and product performance more deeply and is of great significance to product optimization and innovation.

#### **4. Application of Embedded Experiments in Automatic Speech Recognition**

The combination and application of automatic speech recognition (ASR) and embedded experiments show great potential. In embedded systems, the user's voice signal needs to be collected through a microphone first. The collected voice signal usually needs to be preprocessed, including noise suppression, voice enhancement, endpoint detection, etc., to improve the accuracy of voice recognition. The preprocessed voice signal will be converted into a set of feature parameters, such as Mel-frequency cepstral coefficients (MFCC). ASR in embedded systems usually uses HMM or deep neural networks (DNN) for voice recognition. The model-matching process includes a combination of acoustic models, language models and pronunciation dictionaries. By calculating the matching degree between the input voice features and the acoustic model, possible word sequences are obtained. Due to the limited computing power and storage resources of embedded systems, ASR algorithms need to be optimized on embedded devices. Common methods include model compression, pruning, quantization, etc. to reduce computational complexity and memory usage. In addition, hardware

accelerators (such as DSP, FPGA) can be used to improve computing efficiency. To achieve real-time speech recognition in embedded systems, it is necessary to ensure that the latency of the speech recognition algorithm is low enough. This usually requires efficient algorithm design and system optimization. These basic principles constitute the technical foundation for the combination of ASR and embedded systems and promote the widespread application of ASR technology in the Internet of Things, smart home, medical health and industrial fields.

The widespread application of cutting-edge computer technology has revealed the importance of artificial intelligence in almost all economic fields. As one of the core technologies, ASR plays a key role in realizing voice control functions in many Internet of Things (IoT) devices. By combining with embedded experiments, ASR can not only optimize product or system performance through user feedback during use but also expand its application value in multiple fields. A typical case of embedded experiments combined with ASR is to study how to use the optimal transmission data method to apply performance scaling to select ASR data from some disclosed data sources. The study selected two different ASR tasks as experimental objects, evaluated the data selection strategy under the fine-tuning model, and proposed a method to optimize ASR performance [13]. In the field of smart homes, Lecouteux et al. proposed a method for directly integrating multi-command assignment in the ASR system and set up an experiment to verify this method. A space for observing the intelligent interaction between users and the environment was established to obtain a relevant corpus of home automation commands. The experiment collected the SWEET-HOME voice corpus in the DOMUS smart home, which contains home commands, emergency calls and daily sentences, and was recorded using a ceiling microphone. The corpus was used to test and tune two ASR systems, Sphinx and Speeral. The experiment was divided into two stages: daily conversation and reading predetermined sentences. The acoustic model was adapted by MAP and MLLR methods, and linear interpolation of the language model was explored. The results showed that the language model with MLLR adaptation and mild interpolation significantly improved the ASR performance, among which Speeral performed best [14]. The combination of embedded experiments and automatic speech recognition is also reflected in the medical field. Zeng et al. evaluated the application of embedded experiments and ASR in medicine. To assist telemedicine, the researchers developed a "virtual doctor" system that allows AI to form a natural conversation with patients, thereby collecting the patient's condition and medical history, and ultimately providing intervention measures for the patient. In order to realize this system, the researchers need to establish a dataset and test it. The researchers used the Chinese MedDialog dataset to compare the performance of three models, Transformer, BERT-GPT and GPT, in generating medical dialogue responses. The dataset is divided into training set, validation set and test set in a ratio of 0.8:0.1:0.1. The model is constructed at the Chinese character level and trained using its own specific hyperparameters and pre-training data. During the training process, the validation set is used to tune the hyperparameters and the training is stopped when the validation loss no longer decreases. Finally, a variety of automatic evaluation metrics (such as perplexity, NIST, BLEU, METEOR, Entropy, and Dist) are used to evaluate the quality, similarity, and diversity of the generated responses [15].

Although this combination has shown broad application prospects in various fields, it still faces many challenges in practical applications, such as accent recognition, language diversity, computing power and memory limitations in ASR technology. Perhaps this limitation can be further improved in the future through model optimization, model compression, model pruning and multimodal interaction, thereby achieving more intelligent and convenient applications.

## 5. Conclusion

This article reviews the research background and purpose of combining ASR with embedded systems and discusses practical applications in the Internet of Things, smart homes, medical health and other



fields. Research has found that evaluating the application of ASR or its related content in devices or technologies through embedded experiments can significantly improve the intelligence level and user experience of the device. For example, in a smart home environment, ASR technology enables users to control various devices in the home through voice commands, improving convenience and operational efficiency, while applying embedded experiments can collect users' reactions during use, thereby automatic speech recognition in terms of accent recognition, environmental noise, and computing resources. Future research can further improve the performance of the ASR system through model optimization, model compression, hardware acceleration and other methods. In addition, the application of multi-modal interaction technology is also expected to bring new breakthroughs to ASR technology. Overall, the research in this article provides a valuable reference for the combination of ASR and embedded systems, and proposes optimization directions in practical applications, pointing out the way forward for the development of related fields.

## References

- [1] Yu, D., & Deng, L. (2016). *Automatic speech recognition (Vol. 1)*. Berlin: Springer.
- [2] Delić, V., Perić, Z., Sečujski, M., Jakovljević, N., Nikolić, J., Mišković, D., ... & Delić, T. (2019). *Speech technology progress based on new machine learning paradigm. Computational intelligence and neuroscience*, 2019(1), 4368036.
- [3] Bhardwaj, V., Ben Othman, M. T., Kukreja, V., Belkhier, Y., Bajaj, M., Goud, B. S., ... & Hamam, H. (2022). *Automatic speech recognition (asr) systems for children: A systematic literature review. Applied Sciences*, 12(9), 4419.
- [4] Froiz-Míguez, I., Fraga-Lamas, P., & Fernández-Caramés, T. M. (2023). *Design, Implementation, and Practical Evaluation of a Voice Recognition Based IoT Home Automation System for Low-Resource Languages and Resource-Constrained Edge IoT Devices: A System for Galician and Mobile Opportunistic Scenarios. IEEE Access*, 11, 63623-63649.
- [5] Alsalm, A. S., & Javed, M. A. (2024). *Efficient and Secure Data Storage for Future Networks: Review and Future Opportunities. IEEE Access*.
- [6] Kadhim, I. J., Abdulabbas, T. E., Ali, R., Hassoon, A. F., & Premaratne, P. (2024). *A Enhanced Speech Command Recognition using Convolutional Neural Networks. Journal of Engineering and Sustainable Development*, 28(6), 754-761.
- [7] Vadwala, A. Y., Suthar, K. A., Karmakar, Y. A., Pandya, N., & Patel, B. (2017). *Survey paper on different speech recognition algorithm: challenges and techniques. Int J Comput Appl*, 175(1), 31-36.
- [8] Malik, M., Malik, M. K., Mehmood, K., & Makhdoom, I. (2021). *Automatic speech recognition: a survey. Multimedia Tools and Applications*, 80, 9411-9457.
- [9] Jetté, M., & Miller, C. (2022). *The Future of Speech Recognition: Where will we be in 2030?. The Gradient*.
- [10] Gellie, N. J., Breed, M. F., Mortimer, P. E., Harrison, R. D., Xu, J., & Lowe, A. J. (2018). *Networked and embedded scientific experiments will improve restoration outcomes. Frontiers in Ecology and the Environment*, 16(5), 288-294.
- [11] Oliveira, F., Costa, D. G., Assis, F., & Silva, I. (2024). *Internet of Intelligent Things: A convergence of embedded systems, edge computing and machine learning. Internet of Things*, 101153.
- [12] Abdulmalek, S., Nasir, A., Jabbar, W. A., Almuahaya, M. A., Bairagi, A. K., Khan, M. A. M., & Kee, S. H. (2022, October). *IoT-based healthcare-monitoring system towards improving quality of life: A review. In Healthcare (Vol. 10, No. 10, p. 1993). MDPI*.
- [13] Just, H. A., Chen, I. F., Kang, F., Zhang, Y., Sahu, A. K., & Jia, R. (2023). *ASR data selection from multiple sources: A practical approach on performance scaling*.
- [14] Lecouteux, B., Vacher, M., & Portet, F. (2011, May). *Distant speech recognition for home automation: Preliminary experimental results in a smart home. In 2011 6th Conference on Speech Technology and Human-Computer Dialogue (SpeD) (pp. 1-10). IEEE*.
- [15] Zeng, G., Yang, W., Ju, Z., Yang, Y., Wang, S., Zhang, R., ... & Xie, P. (2020, November). *MedDialog: Large-scale medical dialogue datasets. In Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) (pp. 9241-9250)*.