International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 14, Issue 3 [March. 2025] PP: 79-87

Multi-scene gesture recognition system based on machine vision

Binggao He[1], Qianzhang zhuang[1] ,Xuye Zhuang[1] ,Ye Tang[1] ,Zhe Wang[1]
[1] *Shool of Electronic Information Engineering, Changchun University, Changchun, CHINA*
Corresponding Author: Binggao He

ABSTRACT: *The signatory and gesture recognition system based on machine vision technology has important research background and far-reaching significance. In terms of theoretical research, the research and development of gesture recognition system involves the deep crossover and integration of many disciplines such as computer vision, machine learning, pattern recognition and so on. Gesture recognition technology can also be integrated with speech recognition, natural language processing and other technologies. From the perspective of practical application, the gesture recognition system has broad application prospects and great social value. In the field of special education, the system can significantly improve the communication efficiency between deaf students and hearing students. In the field of public service, the application of this system can significantly improve the quality and efficiency of services for deaf and mute people. It also has a wide range of application potential in smart home, robot and other fields.*

Date of Submission: 06-03-2025 Date of acceptance: 18-03-2025

I. PROJECT BACKGROUND AND INTRODUCTION

China has the largest number of hearing disabled people in the world. According to the results of the second national sample survey of disabled people, there are 27.8 million people with hearing disabilities in China (as shown in Figure 1 of the survey chart), this number may be far more than ordinary people imagine, and because of the particularity of this group is relatively less easy to voice in the public vision. Although China's investment in information barrier-free construction is gradually increasing, the communication barriers of hearing-impaired people in many scenes still need to be solved. "How to let them also enjoy the convenience brought by modern science and technology" is not only a technical issue, but also a livelihood issue. Although there are now devices that convert speech into gestures to provide services for people with hearing impairment, communication between people is by no means one-way communication, and how to convert gestures into speech to achieve two-way communication is still a technical problem. Therefore, gesture recognition technology is expected to bring benefits to the hearing impaired.LIU TING TING[1].

Sign language recognition is an important research direction in the field of computer vision, which aims to recognize and understand the sign language gestures used by deaf and mute people through computer technology. Sign language, as a kind of natural language, has rich semantic meaning and expressive power, and is an important means of communication and expression for deaf and mute people. However, due to the limitations of hearing impairment, there are huge barriers for deaf and mute people to communicate with able-bodied people. The appearance of sign language recognition technology provides a new way for the interaction between deaf-mutes and computers, and also provides a more convenient way for the communication between deaf-mutes and healthy peopleWANG TIAN RAN[2].

The purpose of sign language recognition research is to automatically recognize and understand sign language gestures through computer technology, so as to provide more intelligent and efficient human-computer interaction experience for deaf people. The research in this field not only has important theoretical value, but also has wide application prospect.KO XING[3] For example, sign language recognition technology can be applied to virtual reality, smart homes, robots and other fields to provide more convenient and intelligent services for deaf people.

The research of sign language recognition began in the 1980s, and after many years of development, certain results have been achieved. At present, deep learning-based methods have made remarkable progress in the field of sign language recognition. Among them, convolutional neural network (CNN) and recurrent neural network (RNN) are among the most commonly used models. Through the steps of pre-processing, feature extraction and classification of gesture images, these models can realize the automatic recognition of gesture. In addition, the size and quality of the data set is critical to the accuracy of sign language recognition. Currently available sign language datasets include ASL-Letters, ASL-gestures, Tess gestures, etc. These datasets provide researchers with valuable training and testing data.HUANG PU ZHAN LU [4].

Although sign language recognition has achieved certain results, there are still many challenges and problems that need to be solved. Firstly, the variety and complexity of gestures is one of the major challenges facing sign language recognition. There are great differences in sign language gestures in different regions and different cultural backgrounds, which brings great difficulty to the recognition of gestures. In addition, gesture posture, illumination, occlusion and other factors will also affect the recognition accuracy. Secondly, the limitation of data set is also an important problem in sign language recognition. Due to the complexity and diversity of sign language gestures, labeling large amounts of data takes a lot of time and labor costs. Therefore, how to use limited data to train efficient sign language recognition model is an urgent problem to be solvedZHU JUN [5].

II. PRODUCT CORE CONTENT AND TECHNOLOGY

In this project, we plan to build a hardware system based on STM32 MCU, CCD camera module, power management module and WIFI communication module. The main goal of this system is to realize efficient gesture recognition, especially for the capture and recognition of mime gestures, in order to provide a more convenient and intelligent communication tool in mime.

First of all, we choose STM32 MCU, which is a 32-bit microprocessor, with 48 standard pins, 72MHz main frequency, the maximum 20K running memory and 64K flash program memory to provide strong hardware support for our project.

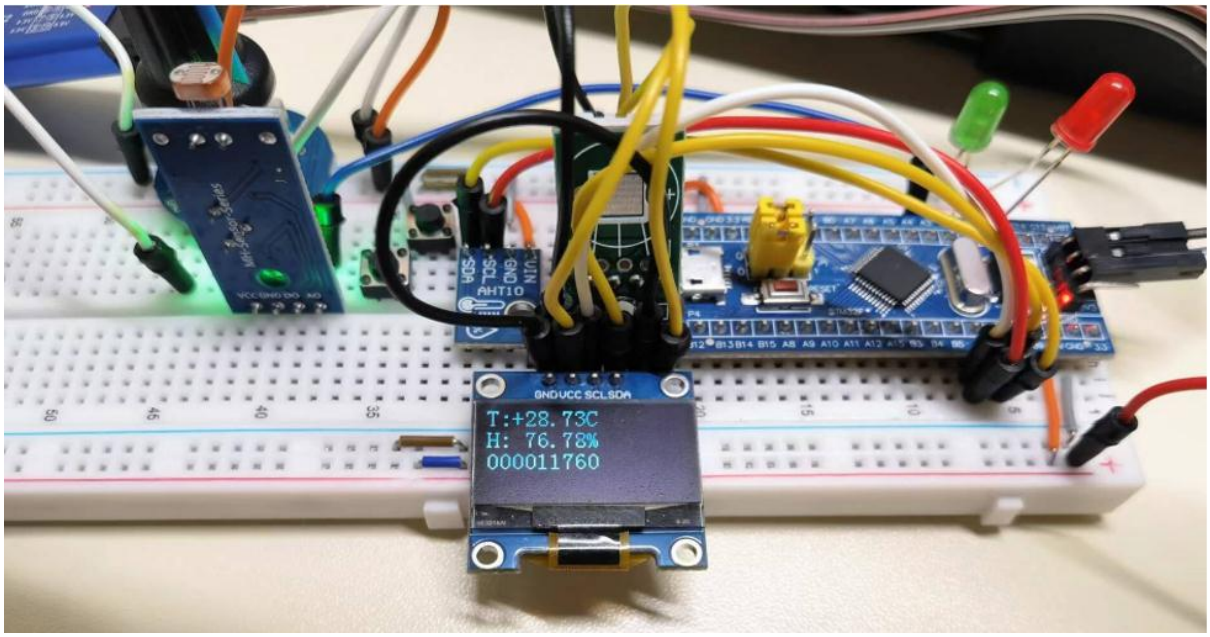

Fig. 1 STM32 MCU development board

Then the CCD camera module will take on the important task of capturing gesture images. This module needs to have the characteristics of high definition, high sensitivity and fast response to ensure that we can capture clear and accurate gesture images. By carefully selecting and optimizing the configuration of the camera module, we can obtain high-quality gesture image data, which provides strong support for the subsequent recognition algorithm.

Fig. 2 CCD HD camera

The power management module is the key to ensure the continuous and stable operation of the system. It will be responsible for providing a stable and reliable power supply for the entire system, ensuring that each module gets enough power under normal working conditions. By using efficient power management technology and rational circuit design, we can reduce the energy consumption of the system,
Improve the overall system performance

Fig. 3 Power management module

WIFI communication module is the key to realize system remote control and data transmission. Through the WIFI communication module, we can connect the system to the Internet and achieve communication with remote servers or other devices. In this way, we can realize the real-time monitoring, data analysis and upgrade maintenance of the system through remote control and data transmission.

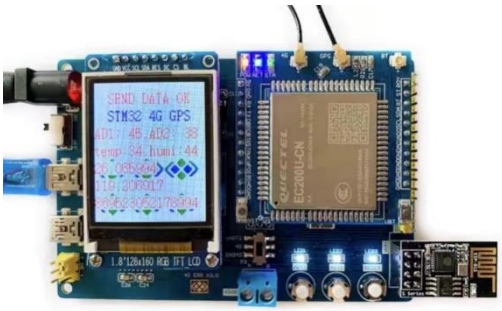

Fig. 4 WIFI communication module

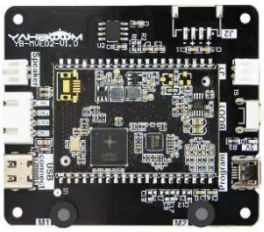

Fig. 5 Voice interaction module

After the hardware system is built, we will build an efficient gesture recognition algorithm. This algorithm needs to be able to accurately capture and recognize gestures, and has high real-time and accuracy. To achieve this, we will employ advanced techniques such as deep learning to train and optimize our algorithms by collecting and processing large amounts of data on gestures. Through continuous experiments and improvements, we can improve the performance and recognition accuracy of the algorithm, and provide users with better services.

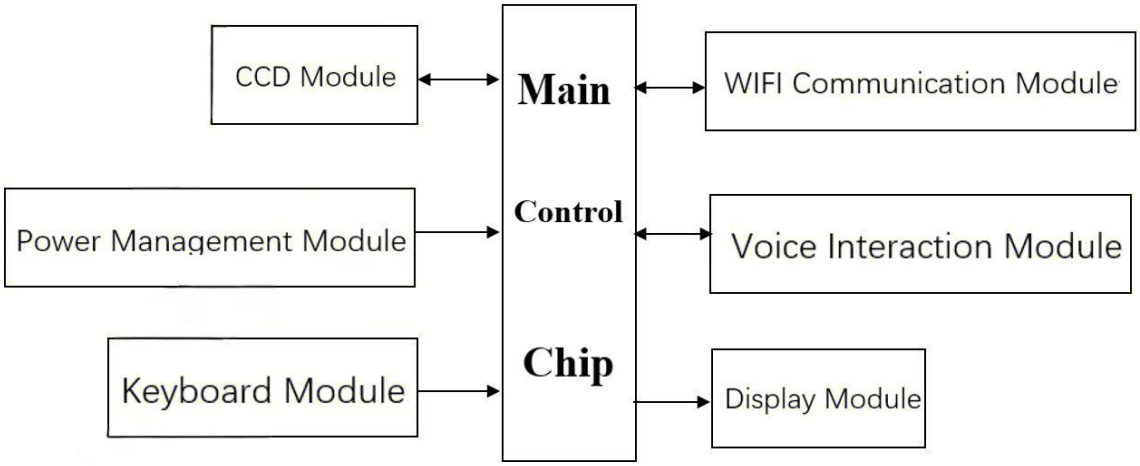

Fig. 6 system schematic chart

At the same time, we will also collect and process diverse data on gestures. The data will include the gestures of different people, gestures in different environments and gestures at different speeds. Through the collection and processing of these data, we can have a more comprehensive understanding of the characteristics and rules of gestures, and further optimize our algorithm and interface design.

Fig. 7 keyboard module

In terms of interface design, we will focus on user experience and friendly interaction methods. We will use simple and clear interface design and intuitive and easy to use interaction to reduce the user's use threshold. At the same time, we will continue to improve and optimize our interface design according to user feedback and needs.mprove user experience and satisfaction.

Finally, in the process of system design and implementation, we will adopt the idea of modular design. Through modular design, we can divide the system into several relatively independent modules, each module is responsible for completing a specific function. This can not only reduce the complexity and development difficulty of the system, but also improve the maintainability and scalability of the system. When the system needs to be upgraded or maintained, we only need to modify and optimize the corresponding modules, and do not need to redesign and develop the entire system.

III. APPLICATION OF THE PRODUCT AND ITS DRAWING

The application of gesture recognition is extensive and diverse, it is not only an important bridge between deaf and deaf people and hearing people, but also plays an important role in many fields. The following is a detailed analysis of the main areas of gesture recognition applications and their specific applications:

Educational field:

Schools for the deaf and special education institutions: Sign language is the main communication tool in these institutions. Through gesture recognition technology, teachers can be assisted in teaching, so that students can better understand and master knowledge. The application of this technology provides a more equal and inclusive learning environment for students and teachers.

Sign language education: Gesture recognition technology makes the process of sign language education more smooth and intuitive. Students can receive information through real-time translation, making the learning process more natural and straightforward.

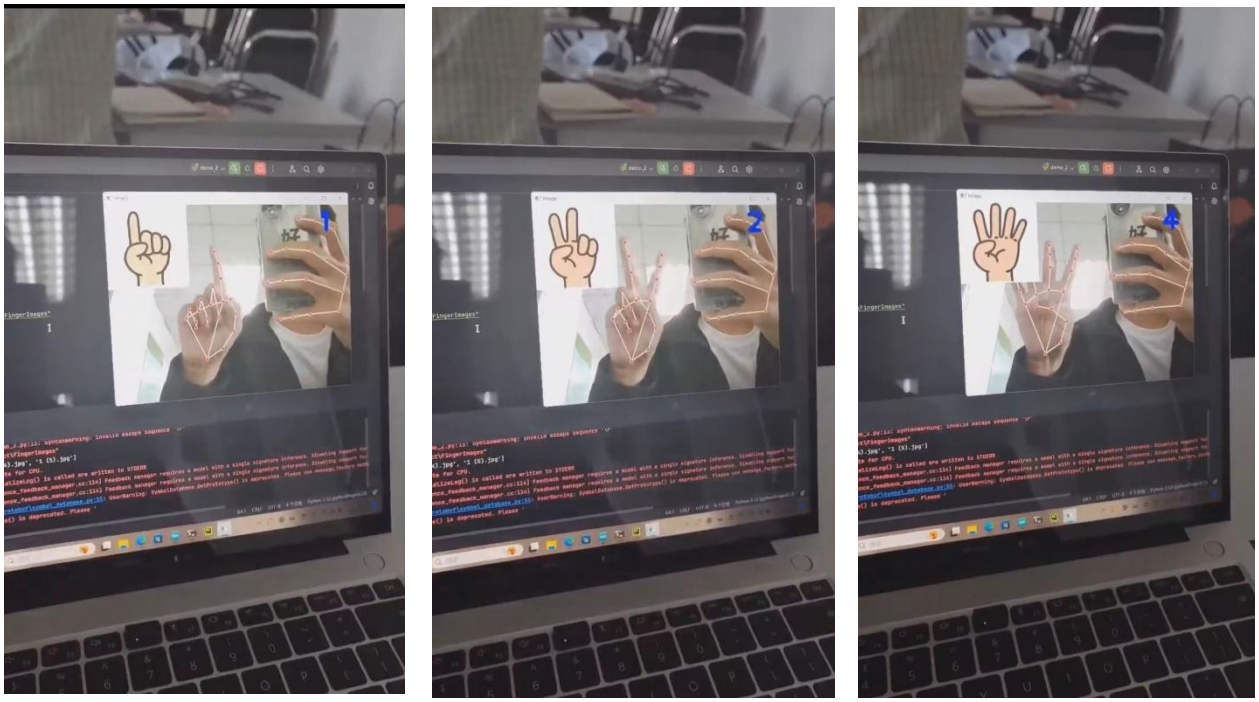

Fig. 8.9.10 LANGUAGE DESTURE RECOGNITION BACKGROUND OPERATION SYSTEM

Medical fields

Communication between doctors and hearing impaired patients: Through gesture recognition technology, doctors can more accurately understand the needs and conditions of patients, thus providing more accurate diagnosis and treatment.

Rehabilitation center: In the rehabilitation center, deaf patients can obtain better rehabilitation services through gesture recognition technology and improve rehabilitation effects.

Social application fields

Sign language translation: Gesture recognition technology can realize real-time conversion between sign language and speech, making communication between deaf people and hearing people easier and more direct. Some real-time sign language translation systems are already able to detect gestures from video content and simultaneously perform speech translation.

Fig. 11 EXAMPLES OF SOCIAL SIGN LANGUAGE APPLICATION

Travel guide

In the tourism industry, sign language and gesture recognition technology can help travel guides communicate more effectively with international tourists, so that tourists can better understand the local culture and attractions.

Emergency rescue

In emergency situations, gestures can help people communicate important information, such as calling for help and asking for help. Gesture recognition technology can help rescue workers understand and respond to these emergency messages more quickly.

Intelligent monitoring

Gesture recognition technology can be used in intelligent monitoring systems to assist security monitoring and early warning by recognizing and analyzing gestures.

Virtusl reality and gaming entertainment

Use gestures to directly control people or objects in a virtual environment, giving users a more immersive experience.

Fig. 12 GESTURE RECOGMITION OF THE FINISHED MAP

To sum up, gesture recognition technology plays an important role in many fields such as education, medical treatment, social communication, tourism, emergency rescue and so on. With the continuous progress of technology and the continuous expansion of application scenarios, gesture recognition technology will bring more convenience and possibility for the communication between deaf people and hearing people.

IV. THE DEVELOPMENTAL STAGES OF GESTURE RECOGNITION

(1) Overall planning and development

The overall planning and development of machine vision multi-scene gesture recognition system is a multi-level, multi-stage complex process, involving technology research and development, scene application, user experience and other aspects. The following is a detailed analysis of its master-planned development:

1.technology development stage

(1) Basic technology research and development: First of all, it is necessary to deeply study core technologies such as machine vision, deep learning, and neural networks to improve the accuracy and stability of gesture recognition. Through continuous optimization of algorithms and models, the precise capture and analysis of hand movements, gestures, expressions and other details are realized.

(2) Multi-scenario adaptability research and development: Develop a targeted gesture recognition system for different application scenarios (such as education, medical treatment, social communication, etc.). This needs to consider the lighting conditions in different scenarios, background noise, user habits and other factors to ensure that the system can run stably in a variety of environments.

(3) Real-time and interactive improvement: optimize the real-time performance of the system, reduce the identification delay, and improve the user experience. At the same time, the interactivity of the system is enhanced, so that users can interact with the system naturally and smoothly through simple gestures.

2. Scene application stage

(1) Application in the field of education: Gesture recognition system can be applied to deaf education, distance education and other fields. Through gesture recognition technology, help deaf students better understand and master knowledge, improve their learning effect. At the same time, it can also provide a more rich and intuitive interactive way for distance education.

(2) Application in the medical field: In the medical field, gesture recognition system can be used to assist diagnosis, rehabilitation treatment and other aspects. For example, doctors can use gesture recognition technology to quickly obtain patient information and develop more accurate treatment plans. At the same time, it can also help patients to carry out rehabilitation training and improve the rehabilitation effect.

(3) Social and entertainment applications: Gesture recognition technology can also be applied to social entertainment to provide users with a more convenient and interesting interactive experience. For example, in virtual reality, augmented reality and other scenarios, users can interact with virtual characters through gestures to enhance immersion and fun.

3. User experience optimization stage

(1) Interface design optimization: According to the user's habits and feedback, optimize the interface design of the system to make it more concise, intuitive and easy to use. At the same time, provide personalized setting options to meet the needs of different users.

(2) Privacy and security: Strengthen the privacy protection and security performance of the system to ensure the security and privacy of user data. Adopt encryption technology, rights management and other measures to prevent data leakage and abuse.

(3) Continuous iteration and upgrade: According to user feedback and market changes, continue to iterate and upgrade the system, and constantly improve its performance and functions. At the same time, pay attention to the development and application of new technologies, and introduce the latest technological achievements into the system.

In short, the overall planning and development of the multi-scene mime gesture recognition system of machine vision needs to consider multiple aspects such as technology research and development, scene application, and user experience. Through continuous optimization and innovation, the rapid development and wide application of the system can be promoted, and a more convenient and efficient communication experience can be brought to deaf people and more users.

(2) Phased development plan

With the rapid development of artificial intelligence technology, machine vision is more and more widely used in the field of gesture recognition. This paper aims to propose a phased development plan to provide clear guidance for the development and implementation of multi-scene gesture recognition system for machine vision. The program will focus on technology research and development, hardware device selection and optimization, data set construction and annotation, algorithm model training and testing, system integration and debugging, multi-scenario application testing, user feedback and improvement, and continuous optimization and iteration.

1. Technology research and development

(1) In-depth study of core technologies such as machine vision, deep learning, and neural networks to improve the accuracy and stability of gesture recognition.

(2) In view of the complexity and diversity of gestures, research effective feature extraction and recognition algorithms.

(3) Explore new technical methods and means to improve the real-time and interactive nature of the system.

2. Hardware equipment selection and optimization

(1) Select the appropriate image acquisition equipment and processor according to the system requirements and performance requirements.

(2) Optimize the hardware equipment to ensure its stability and reliability in various scenarios.

(3) Research new hardware equipment to improve the overall performance and efficiency of the system.

3. Data set construction and annotation

(1) Collect a large amount of data of gestures, including different scenes, different gestures, different characters, etc.

(2) Pre-process and label the data to ensure the quality and accuracy of the data.

(3) Construct standardized data sets to provide strong support for the training and testing of algorithm models.

4. Algorithm model training and testing

(1) The constructed data set is used to train the algorithm model and optimize the model parameters and structure.

(2) Conduct performance testing and evaluation of the model to ensure its accuracy and stability in practical applications.

(3) Constantly improve and optimize the algorithm model to improve the recognition effect.

5. System integration and debugging

(1) Integrate the algorithm model with the hardware device to build a complete gesture recognition system.

(2) Debug and optimize the system to ensure its stability and reliability in various scenarios.

(3) Develop system interfaces and communication protocols to achieve seamless docking with other systems.

6. **Multi-scenario application test**

(1) Conduct systematic application tests in different scenarios, including education, medical treatment, social interaction, etc.

(2) Collect user feedback and data, analyze system performance and user experience.

(3) Improve and optimize the system according to the test results and user feedback.

7. **User feedback and improvement**

(1) Establish a user feedback mechanism and actively collect users' opinions and suggestions on the system.

(2) Localization of problems and formulation of improvement measures based on user feedback.

(3) Timely repair system loopholes and defects to improve user satisfaction.

8. Continuous optimization and iteration

(1) Continue to pay attention to industry dynamics and technological development, and constantly update and optimize system functions and performance.

(2) Maintain and upgrade the system regularly to ensure the stability and security of the system.

Constantly expand the application scenarios and scope of the system, and promote the wide application and development of gesture recognition technology.

Ⅴ.CONCLUSION

The multi-scene gesture recognition technology based on machine vision will show an extremely broad development prospect and profound application potential. With the vigorous development of artificial intelligence technology, especially the continuous breakthroughs in the fields of deep learning, computer vision and pattern recognition, the performance of gesture recognition systems is rapidly improving. This technology can not only greatly promote the accessibility of communication with deaf people, but also play an important role in many fields such as education, health care, services and so on.

With the continuous progress of technology and system optimization, we can expect that the recognition accuracy of gesture recognition technology will be higher and higher, reaching or even exceeding the human level. This means that the system will be able to more accurately understand the gesture intentions of deaf people, thus reducing misunderstandings and communication barriers. At the same time, the improvement of recognition speed will make the communication more smooth, close to the speed of natural dialogue, so that deaf people enjoy the experience of equal communication with non-deaf people in real-time communication.

In addition, the future of gesture recognition technology will also be combined with other advanced technologies, such as natural language processing (NLP) and speech recognition technology, to create a multi-modal communication platform. With the help of NLP, the system can better understand the context and meaning behind gestures, and provide more accurate language translation and expression assistance for deaf and mute people. Combined with speech recognition, deaf people can communicate with people with normal hearing

through speech, thus breaking the boundaries of communication and realizing two-way communication in a true sense.

With the evolution of technology, it is expected that this multi-scene gesture recognition system will become more intelligent and personalized. The system may use machine learning to analyze the user's communication habits and preferences to provide customized services. For example, in a specific industry environment, the system can optimize the recognition algorithm based on industry terminology to improve the recognition accuracy in a specific context.

The future of gesture recognition technology will also be widely used in wearable devices. For example, devices such as smartwatches or gloves can monitor and interpret gestures in real time, providing a more convenient way for deaf people to interact. This integrated solution can not only help deaf people communicate more easily in their daily lives, but also play a huge role in professional fields such as sign language interpretation, teaching and coaching.

In addition, with the development of cloud computing and Internet of Things technology, the gesture recognition system is expected to be closely integrated with concepts such as smart home and smart city, creating more innovative application scenarios. For example, deaf people can use gestures to control smart devices in their homes, such as adjusting lights, temperature, etc., thereby improving their quality of life and convenience.

To sum up, multi-scene gesture recognition technology based on machine vision will have unlimited possibilities in the future. With the deepening of technological innovation and interdisciplinary integration, we have reason to believe that this technology will gradually break down communication barriers, bring more hope and opportunities to deaf people, and promote the development of human society towards a more inclusive and barrier-free direction.

REFRENCES

[1]. Liu Tingting. Research on Sign language gesture recognition of special population based on millimeter wave radar, Master of Northwest Normal University 2024-06-30.

[2]. WANG Tanzi, Wang Qi, Wang Qingshan. Cross-object sign language gesture recognition method based on transfer learning, Computer Science, 2023-06-15.

[3]. Kou Xin. Research on Cross-domain Sign Language Gesture Recognition based on millimeter wave signal, Master of Northwest University 2022-06-30.

[4]. Huang Fu Zhanlu. Dynamic Mime Gesture Recognition based on lightweight deep learning network, Master of Nanchang University 2022-05-28.

[5]. Zhu Jun. Research on gesture recognition Algorithm based on Deep learning, North China University of Technology 2022-05-15 Master.