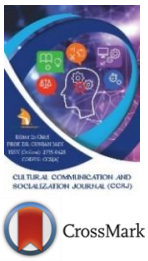
ZIBELINE INTERNATIONAL™ PUBLISHING

Cultural Communication and Socialization Journal (CCSJ)

ISSN: 2735-0428 (Online)
CODEN: CCSJAJ

DOI: http://doi.org/10.26480/ccsj.02.2023.66.69



REVIEW ARTICLE

BASED ON THE HELP AND DEVELOPMENT ANALYSIS OF YOLOv5s SIGN LANGUAGE INTERPRETER FOR THE HEARING IMPAIRED

Xiaoyu Liu, Yifan Zhang, Haowen He, Diwei Shi, Lin Long*

Electrical Engineering, Southwest MinZu University, Chengdu, China
****Corresponding Author E-mail: 33660443@qq.com***

ARTICLE DETAILS

Article History:

Received 23 May 2023
Revised 06 June 2023
Accepted 10 July 2023
Available online 13 July 2023

ABSTRACT

In order to help the hearing-impaired communicate with others in social life, promote the employment of the disabled, and do our best to improve the quality of life of the hearing-impaired, this project proposes a self-training hand model based on YOLOv5s, and adds CBAM attention mechanism to improve it, so as to improve the model recognition accuracy of the sign language interpretation machine to a certain extent. In view of the practical problem that the assistive tools that the hearing impaired people can choose in the market have certain limitations, this paper mainly introduces the basic structure and innovation points of the sign language interpreter researched by the project, the main application technology, the specific help for the hearing impaired and the development prospect of the sign language interpreter. Therefore, it is concluded that the sign language interpreter studied in this project has certain development prospects. The experimental results show that the average accuracy of the improved algorithm is 97.01%, which is 2.78 percentage points higher than that of the original YOLOv5s, and the number of parameters is not much different, which can meet the daily life Demand.

KEYWORDS

sign language to text, YOLOv5s, barrier-free communication, Sign language interpreter

1. INTRODUCTION

We need a lot of communication and communication in our lives, but in most cases, hearing-impaired people will cause a lot of inconvenience due to not fully understanding the expression of the communicator, so that the circle of communication continues to shrink and form depression and anxiety. According to statistics, China is the country with the largest number of hearing disabilities in the world, and there are about 27.8 million hearing disabilities, accounting for about 6% of the total number of hearing disabilities in the world (Yang et al., 2022). Combined with data released by the World Health Organization in 2018, more than 466 million people worldwide suffer from disabling hearing loss. People with hearing impairments make up a large proportion of the disabled population worldwide, and as people learn more about special groups, sign language is paid more attention (Abula et al., 2021).

With the passage of time, the aging of the global society population is becoming more and more common, and hearing impairment is a major public health problem facing the world. Taking this as a starting point, the project team researched and developed a sign language interpretation machine specially designed to solve the inconvenience of the hearing impaired. The sign language interpretation machine has the characteristics of easy carrying, fast sign language translation, more accurate sign language interpretation and text reading function. The realization of this sign language interpreter makes up for the shortage of assistive tools used by the hearing impaired, and can greatly improve the efficiency and quality of the daily life and work of the hearing impaired.

Single-stage object detection algorithms such as YOLO series generally have the advantage of fast detection speed, although the accuracy is lower than that of two-stage object detection algorithms, but after improvement, the accuracy also fully meets the requirements of real-time detection. The YOLO series itself is also constantly improving and developing, and the effect in accuracy and detection speed is getting better and better. Compared with YOLOv4, YOLOv5 is lightweight and has great flexibility, which can well meet the requirements of real-time sign language image detection.

The sign language real-time interpretation system based on YOLOv5s proposed in this paper trains the model for sign language interpretation to achieve a predetermined standard accuracy, and then puts the model into the relevant hardware. The model has less weight, fast translation, and fits well into embedded devices. The system can translate the sign language displayed in front of the screen in real time, and the speed and accuracy are relatively high.

Sign language recognition refers to the use of algorithms and technology to recognize the resulting sequence of gestures and explain their meaning in text or speech. This project aims to investigate a new device that can help the hearing-impaired improve their quality of life, a sign language interpreter based on the YOLOv5s self-training hand model and CBAM attention mechanism to improve the accuracy of sign language recognition. At present, it has made its own sign language data set, and the sign language interpreter has been able to successfully identify and carry out Chinese voice broadcast, and the relevant functions have been basically realized, and the specific details are still being improved. This article will share the basic structure, innovation points, application technology and functions of sign language interpreters in turn, and finally look forward to the development trend of the device.

2. CURRENT STATUS OF ASSISTIVE DEVICES USED BY HEARING-IMPAIRED PERSONS

At present, the assistive tools available for the hearing impaired on the

Quick Response Code	Access this article online	
	Website: www.ccsj.com.my	**DOI:** 10.26480/ccsj.02.2023.66.69

Cite The Article: Xiaoyu Liu, Yifan Zhang, Haowen He, Diwei Shi, Lin Long (2023). Based on the Help and Development Analysis of YOLOv5s Sign Language Interpreter for The Hearing Impaired. *Cultural Communication and Socialization Journal, 4(2): 66-69.*

market mainly include sign language, hearing aids and sign language interpretation service centers.

The direct use of sign language to communicate is the most limited, most ordinary people have never learned sign language in daily life, which makes the gradual formation of an invisible barrier between normal people and hearing impaired people. In addition, most hearing-impaired people only use simple and common sign language to communicate with their families.

Hearing aids are essentially electroacoustic amplifiers, can only help mild to moderate hearing loss users, but most hearing impaired people also suffer from language impairment, hearing aids cannot fully help hearing impaired people communicate with others normally, according to the "China Disability Watch Report (2018)" shows that the biggest problem of cochlear implants in China is the low coverage of beneficiary populations. The reason for this is that there are few cochlear implant brands, high prices, certain requirements for the physiological basis of implanters, which are not covered by medical insurance, and commercial insurance does not support.

In recent years, the sign language interpretation service center has gradually developed from the initial public welfare volunteer service related institutions to enterprises and institutions that meet market demand, but the industry as a whole is in its infancy and has not yet formed a mature market mechanism. Statistics show that only 25.4% of deaf people have used sign language interpretation services, of which only 36.9%[Error! Reference source not found.] (Ni and He, 2022) are satisfied with the services. Service occasions, quality and price are the three most urgent needs of deaf people that need to be improved. Although it can effectively help the hearing impaired solve problems, it cannot always accompany the life of the hearing impaired.

3. The Basic Functional Structure of Sign Language Translator and its Innovation

3.1 Basic Functional Structure

The structure of the sign language interpreter can be divided into two parts: client and cloud computing server, which can realize relatively simple and commonly used gesture collection and translation. The hearing-impaired use sign language in the camera recognition range frame, and the client collects sign language information with the built-in camera and uses the YOLOv5s' self-trained hand model for data processing and conversion, generates the corresponding text, and then reproduces it on the client display, and can also be played out through speech synthesis technology if necessary.

3.2 The innovation of Sign Language Interpreters

3.2.1 It Starts from Public Welfare

The original intention of this project is to respond to the theme of national science and technology for good and help promote the barrier-free construction of today's society. The initial plan of this machine is to carry out public welfare promotion trial operation by putting it into hearing-impaired schools or volunteer organizations. According to the feedback of the effect, various volunteer activities such as volunteer service activities, public fundraising activities or charity sales are held in a timely manner to achieve the dual purpose of project development and brand promotion.

3.2.2 Combine Cloud Technology

The relatively broad and unlimited future development prospects of cloud technology are the main reasons for choosing Baidu AI cloud data analysis for this project. Cloud technology helps this sign language translation machine to achieve faster and more accurate conversion between text and speech. At the same time, considering the "cloud"-centric design, the expansion of computers in the future will allow more and more organizational designs to expand to different industries, and the increase in the number of mobile devices will also make more and more platforms become mobile cloud services.

3.2.3 The goal is to fully integrate the hearing-impaired into social life

This project focuses on people with hearing impairments, and so far there are more than 27 million people with hearing impairment in our country. For people with hearing impairment, in addition to using sign language to communicate, they are almost no different from normal people, they are easily ignored by today's society, cannot timely and autonomously perceive non-verbal signals, such as alarm clocks, doorbells, signal alarms, car horns, mobile phone ringtones, etc., dangers may occur at any time in their lives. In addition, most of them live in a closed self-world, and their motivation to participate in social activities is extremely low. This project can help the hearing-impaired to better integrate into the collective life of society, basically achieve barrier-free communication with others, and be able to express the hearing-impaired people's own ideas in a timely and accurate manner. Starting from adapting to normal social life, gradually mastering the lifestyle of ordinary people, and then starting their own business, so as to help the disabled to achieve self-sufficiency, hoping to make their own contributions to promoting the development of the cause of the disabled.

4. Sign Language Interpreter Image Processing Technology and its Practical Application and Help

4.1 Image processing technology of sign language interpreter

YOLO (You Only Look Once) is one of the popular object detection models and is widely used in the industry. The basic principle of YOLO is: first divide the input image into 7×7 grids, predict 2 borders for each mesh, then remove the target window with low probability according to the threshold, and finally use the border merging method to remove the redundant window to obtain the detection result (Liu et al., 2022). YOLOv5 uses the same cross-stage partial network (CSP) structure as YOLOv4, but YOLOv4 uses only CSP structures in Backbone, while YOLOv5 uses two different CSPs in Backbone and Neck. In this paper, the YOLOv5s network model is selected to improve. The structure of YOLOv5s is divided into four parts: Input (input), Backbone (backbone network), Neck (multi-scale feature fusion module), and Prediction (prediction end), and its network structure is shown in Figure 1 (Xing and Pan, 2022). At present, commonly used Neck aggregate blocks are FPN (feature pyramid networks)[Error! Reference source not found.], PAN (pyramid attention network), ASFF (adaptively spatial feature fusion), BiFPN and so on (Tan et al., 2020).

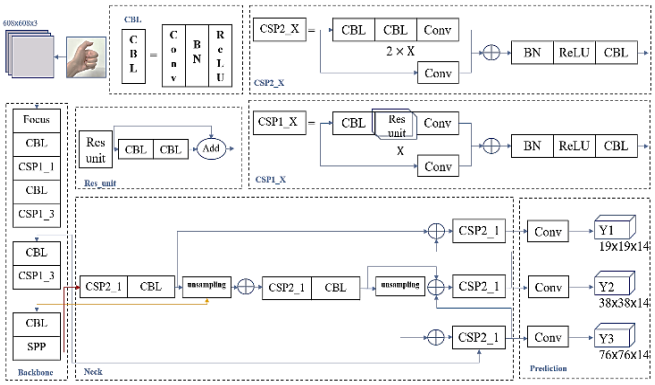


Figure 1: Network structure diagram of YOLOv5s

For the input sign language image, in addition to the hand information, it is often accompanied by complex background information, and when performing convolution operations, the iterative accumulation of this background will form a large amount of unwanted messy information, thereby flooding some targets and resulting in low accuracy (Wang et al, 2021). To this end, this project improves the CBAM attention mechanism, selects effective locations and adds them to the YOLOv5s network model for feature fusion, so that the model can locate and identify sign language expression more accurately.

The CBAM (Convolutional Block Attention Module) attention mechanism takes into account not only the different importance of different feature channels, but also the importance of different locations of the same feature channel. First through the attention mechanism module, and then through the spatial attention module, the effect is slightly better than the way the space and then the channel and the spatial attention module are parallel. The principle of CBAM is shown in Figure 2 (Niu et al., 2021).

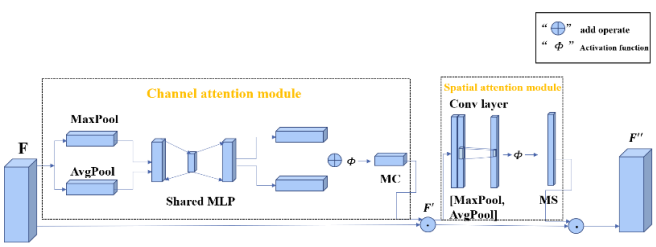


Figure 2: Network structure diagram of CBAM

The results of the two different attention mechanisms used in this project to improve YOLOv5s with CBAM and YOLOv5s alone are shown in Table 1.

Table 1: Comparison of results by different attention mechanisms

Model	Model parameters/10^6	P/%	R/%	Map/50%
YOLOv5s	7.13	88.16	91.31	94.23
YOLOv5s_CBAM	7.30	89.14	91.79	97.01

Analysis of the data in Table 1 shows that when the CBAM module is added to the YOLOv5s network, the model parameters increase by 1.7×105 and the Map increase by 1.66 percentage points, which means that adding spatial attention to the attention mechanism module is more conducive to sign language recognition.

In summary, the application technology of sign language interpreter can be summarized as: adding CBAM mechanism to the popular YOLOv5s self-training hand model for improvement, while improving the algorithm coding format, making sign language datasets by itself, and selecting the more reliable NVIDIA development kit itetson TX2 for research and development, which can greatly improve the recognition accuracy of the sign language interpreter, and finally successfully recognize and carry out Chinese voice broadcast (Zhang, 2019).

4.2 Practical use and help of sign language interpreters

The functions currently developed can meet the two daily needs of ordinary hearing-impaired people.

4.2.1 Communicate freely

Under normal circumstances, according to image recognition, speech recognition and other technologies, in daily communication with people, speech and text content can be quickly translated to and frozen, and gestures are captured and recognized to convert them into text or speech, so as to promote "zero" delay barrier-free communication.

4.2.2 Office meetings

In scenarios such as meetings and multi-person communication, speech recognition technology is used to convert speech into text and accurately present it on the screen, allowing hearing-impaired people to quickly obtain information. At the same time, when the hearing impaired express their own opinions, they can convert the edited text content into voice playback, effectively solving the work difficulties of the hearing impaired.

Perhaps in the eyes of ordinary people, hearing-impaired people look the same as normal people, but their world is silent, and the lack of sound perception makes hearing-impaired people face great inconveniences and challenges in life, work and study. As an important device to improve the hearing environment, hearing aids also have many inconveniences. The sign language interpreter machine studied in the project can provide some help to the lives of the hearing impaired. Make it easier for the hearing impaired to integrate into social life and bring different possibilities to their world.

5. Development Prospect of Sign Language Interpreters

With the rapid development of artificial intelligence technology, scientific and technological personnel in various countries continue to help "science and technology for good", and the tools and technologies available for the hearing impaired continue to develop. Some of the great tools are listed below.

"Voice of the Hand" APP. The "Voice of the Hand" APP has the function of voice and text translation, and provides two major functions: online or offline artificial sign language interpretation services. It is a mobile application that enables barrier-free communication between the hearing impaired and the hearing impressed. It can meet the hearing-impaired people's communication anytime, anywhere, communication in temporary scenarios, and three different needs of deaf sign language communication

"Sound Book" APP. A communication software specially developed for the hearing impaired, which includes two modules of barrier-free communication and language training, the company uses voice recognition technology to help the hearing impaired obtain the voice of people within 4 meters, TV, computer, audio equipment and other speakers. The software is committed to helping the hearing-impaired communicate with others through artificial intelligence technology, and strive to integrate the hearing-impaired into society and build a better life.

"AI Sign Language Interpreter". In 2019, Tencent developed an electronic device that can quickly convert sign language into text. Using SLR, with the help of self-developed computer algorithms, it can automatically distinguish a lot of actions and meanings in sign language expression, which is the largest Chinese sign language recognition dataset in China.

In recent years, under the background of the state's emphasis on accelerating the development of the cause of the disabled, it has united and led the disabled and the people of the whole country to actively participate in the great practice of building a modern socialist country in an all-round way. To build a happy and beautiful life together, domestic universities and science and technology enterprises have begun to pay attention to the projects to help the elderly and the disabled. However, there are still few fully ubiquitous and efficient assistive tools.

The sign language interpreter machine studied in this project uses the NVIDIA development kit itetson TX2 as the client body, the audience is moderately and severely hearing impaired, which can provide them with real-time sign language interpretation, text and speech conversion and other main services, which can help the hearing impaired smoothly communicate with more people, do things conveniently, and make life more convenient, so that they can better integrate into social life in a new way. At the same time, the development of the sign language interpreter function in this study is limited by time, and the current available functions can only basically meet the basic needs of sign language interpretation.

With the progress of the Internet and information technology towards the goal of more broadband, the popularity of mobile terminal equipment is very fast, gradually penetrating into all aspects of people's daily life people's growing demand for emotional communication, hearing impaired people's demand for sign language interpretation has become not only a simple neutral tone output, but also more emotional expression (Pan, 2019; Gu et al, 2022).

6. Conclusion

This project researches and develops a Chinese sign language interpreter based on deep learning, aiming to provide more convenient and accurate sign language recognition and translation services for deaf and mute people. By training and optimizing diversified sign language data, the team members proposed a YOLO algorithm with the improvement of CBAM attention mechanism, which was applied to sign language recognition tasks. In experiments, the system demonstrated higher sign language recognition accuracy and faster processing speed, while being able to convert sign language into corresponding Chinese text output.

In general, a new and efficient solution for Chinese sign language interpreter is proposed, and good experimental results have been obtained, which has a wide application prospect. However, the system still faces some challenges and limitations in practical applications, such as sign language diversity and scene complexity, which need to be further improved and optimized to meet the needs of different users. The research direction of this project makes up for the lack and deficiency of some markets, and the development prospects are huge.

Acknowledgment

This work was supported in part by the College of Electrical Engineering, Southwest MinZu University' Students Innovation and Entrepreneurship Training Program (Project No: 202210656005).

References

Abula, M.N.W.E., Kurban, A.L.P., Xie, Q.N., & Geng, L.T. 2021. A review of sign language recognition methods and techniques. Computer Engineering and Applications. (18), 1-12.

Gu, Y., Xu, Q., Mao, B.S., & Lin, Q.M. 2022. Design of a sign language-to-emotion speech tool to assist the hearing-impaired. Computer Literacy and Technology. (15), 72-74+87. doi:10.14004/j.cnki.ckt.2022.0947.

Liu, J.X., Zhou, X., Zhang, S.F., Zeng, L.H., Duan, L.Z., & Shen, L.S. 2022. Implementation of sign language interpretation system based on artificial intelligence. Scientific and Technological Innovation and Application. (23), 41-43+48. doi:10.19981/j.CN23-1581/G3.2022.23.010.

Ni, L., & He Z.Q. 2022. Survey on the demand and current situation of sign language interpretation services in Shanghai. Chinese Trasition. (04), 113-119.

Niu, R.H., Yang, J., Xing, L.X., & Wu, R.B. 2021. Micro-expression recognition algorithm based on convolutional attention module and dual-channel network. Computer. (09), 2552-2559.

Pan, H. 2019. Design and implementation of intelligent distribution system based on WeChat mini program. Microcomputer Applications. (07), 31-33.

Tan, M., Pang, R. & Le, Q.V. 2020. EfficientDet: Scalable and efficient object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. doi:10.1109/CVPR42600.2020.01079.

Wang, P.F., Huang, H.M., Wang, M.Q., & Le, B.J. 2021. YOLOv5s-FCG : An Improved YOLOv5 Method for Inspecting Riders' Helmet Wearing. Journal of Physics: Conference Series(1). doi:10.1088/1742-6596/2024/1/012059.

Xing, J.C., & Pan, G.Z. 2022. Research on the sign language recognition algorithm of improved YOLOv5s. Computer Engineering and Applications. (16), 194-203.

Yang, S.Y., Tian D., Guo, Y.Y., & Zhao, M. 2022. End-to-end sign language translation research. Journal of Systems Science. (01), 71-77.

Zhang, Y.H. 2019. Minimal application hardware design for the Jetson TX2 platform. Microcontroller and embedded system applications. (10), 52-54.

.