ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

From Gesture to Voice: Advancing Human-Computer Interaction for the Deaf Community

Zichong Lu[1,a,*]

[1]*The Hong Kong Polytechnic University, Hong Kong SAR, China*
[a]*luzichong@126.com*
**Corresponding author*

Abstract: *This essay explores the future of interactive technology with a focus on sign language recognition as a vital means of communication for the deaf and hearing-impaired communities. It outlines the current state of sign language technologies, their limitations, and the significant challenges in achieving seamless human-computer interaction. By reviewing existing innovations such as smart sign language interpreter devices and AI-driven gesture recognition, the paper emphasises the transformative potential of these tools in everyday life, education, and public services. Moreover, it highlights the socio-economic impact of accessible communication technologies and calls for further interdisciplinary research and development. Through enhanced datasets, multimodal models, and ethical integration of artificial intelligence, the future holds the promise of more inclusive and human-centred interaction design.*

Keywords: *Sign language recognition; Human-computer interaction (HCI); Accessibility technology; Deaf communication; Artificial intelligence in assistive tech*

1. Current status of deaf and sign language recognition

In recent years, the problems caused by an ageing population have become apparent, and the existence of a large number of elderly people in need of care and disabled people has made the issue of helping the elderly and disabled a very important and difficult one for society. The large number of people in need of care not only places a great deal of emotional and financial pressure on families, but also places a heavy burden on the government. At the same time, the rapid development of the robotics industry has put forward new requirements for human-robot interaction, and the first stage in the history of human-robot interaction - the interface stage - is no longer able to meet the various needs of people's lives. Natural Language is eagerly awaited[8]. Using the dividends generated by the computer industry to address the problems of an ageing population, this initiative will not only improve the quality of life of people, but also provide a viable solution for caring for people with disabilities.

And of all the special populations, the deaf and disabled groups are the ones most often found in HCI research. However, the paper points out that the limitations suffered by the deaf community are often underestimated by both the general public and researchers, and that their problems remain unresolved despite the fact that technologies like sign language recognition are now gradually being developed. It is the desire of every deaf person to be able to interact with the outside world as easily and freely as normal people.

Sign language is the language used by deaf people, and it is a relatively stable system of expression consisting of hand movements supplemented by facial gestures, a special language that is communicated through movement/visuality .[11]

In addition to natural language (spoken and written), human language (expressions, gestures and hand signals) is also one of the basic ways of human interaction. Compared to human interaction, human-computer interaction is much more dull, so the study of human language understanding, i.e. human language perception, and the integration of human language and natural language information is of great significance to improve human language understanding in computers and to enhance the practicality of human-computer interfaces. As Kyle (1988) states, sign language is a very important part of human language, it is the most informative human language, it has the same expressive power as natural language such as speech and written language, and therefore sign language can be used as a tool in human-computer interaction, and it has a strong visual effect, it is vivid, visual and intuitive[12].

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

The study of sign language not only helps to improve the living, learning and working conditions of deaf people, but also provides them with better services. It can also be used in computer-aided dumb language teaching, bilingual television programmes, virtual human research, special effects in film production, animation, medical research, games and entertainment, and many other areas. In addition, the study of sign language involves many disciplines such as mathematics, computer graphics, robotic kinematics and medicine. Therefore, the study and recognition of sign language is a very interesting topic.

2. Existing technologies

And some empathetic designers have taken the first step towards accessible sign language communication, such as the Smart Sign Language Interpreter bracelet designed by the Prof team, which enables deaf people to use sign language to communicate with hearing people. When a deaf person communicates using sign language, a distance sensor at the tip of the fingernail enables the gesture to be tracked and the smart sign language interpreter bracelet converts it into sound or text, and conversely, a normal person's voice can be converted into text and displayed on the bracelet's screen so that the deaf person can see it (See Figure 1).

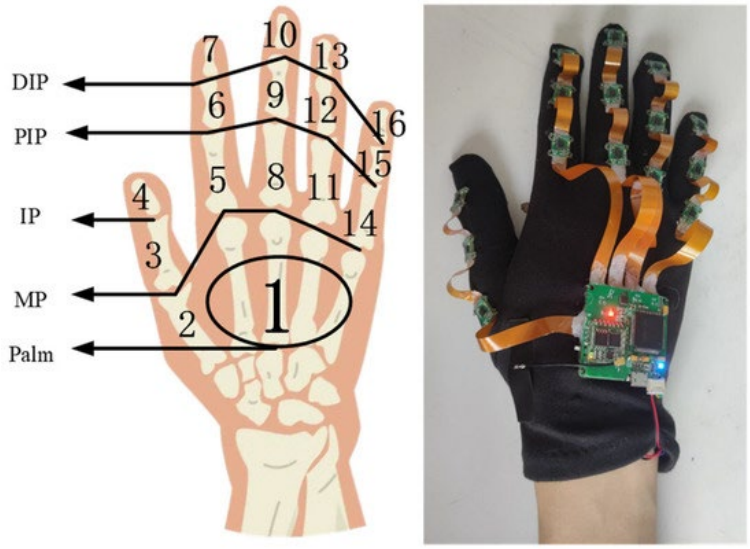

Figure 1:Smart Sign-Language Interpreter Glove: Sensor Placement on Finger Joints and Prototype Hardware

And another example is called signal Glove which is a low-cost wearable translator glove for Mexican Sign Language that won the Mexican national award of the 2024 James Dyson Award (Hernández Jiménez, 2024). The system embeds five fingertip sensors—3D-printed exoskeletal housings with potentiometers and gyroscopes—on each finger joint, and a microcontroller worn on the forearm streams real-time sensor data via Bluetooth to a receiver module with an LED display (See Figure 2). There, a decoding algorithm translates gestures into spelled-out letters on screen, enabling seamless communication between signers and non-signers. Its minimalist 3D-printed design ensures affordability and ease of mass production, while the fully open-source hardware and software framework supports community-driven customization and localization.

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

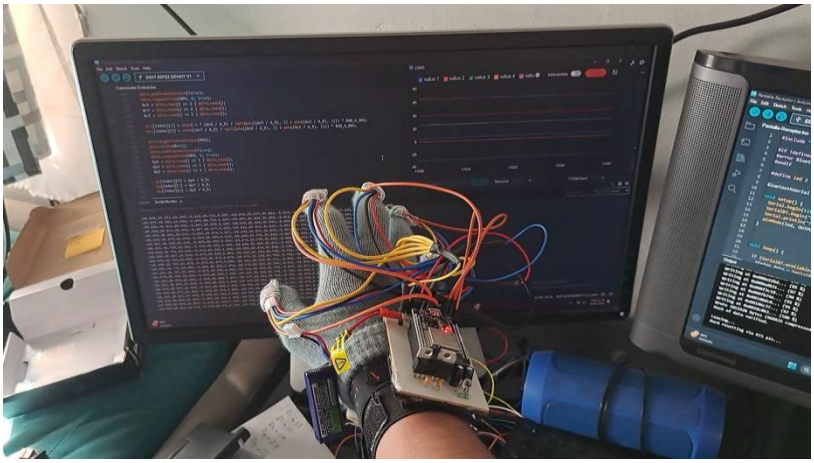

Figure 2:Prototype Wearable Sign-Language Glove with Wired Sensor Array and Real-Time Gesture Processing Interface

3. Shortcomings of current technology

However, due to technical limitations or other factors, most of these technologies are still at the design level and it will take a particularly long time before they are actually brought to the general deaf population and accurately identified.

This is because the deaf community has a very different way of communicating from the general population, and this is reflected in the language they use - sign language. Two concepts have to be mentioned here, vocal language and sign language.

The earliest attempts to bring the deaf community into the digital world originated with two guides, the 1999 WCAG1.0 guide and the 2008 WCAG2.0 guide, both of which had limitations.

Rømen (2012) argues that the 1999 WCAG 1.0 guidelines deal primarily with the tagging and transcription of audio content, but ignore alternative methods related to sign language, with the only reference to sign-language-related content being video management. That is, in practice, captions are added to audio content and, if possible, translated into Sign Languages.These guidelines, although specified in good faith to address the issue of accessibility for deaf people, may have some misconception that written transcriptions of audio content are sufficient to support deaf users.[6] In fact, however, text is closely linked to language, and text-dependent language is a completely different structure of audible language to the sign language that Deaf people were exposed to growing up. So for deaf people, they are not able to read in what appears to us to be a normal language sequence.

Brajnik (2010) mentions that the WCAG 2.0 guidelines, which were revised in 2008, represent a gradual increase in awareness of the needs of deaf people and better address these issues. For example, one of them states that "all pre-recorded audio content provides sign language interpretation in the form of synchronised media types"[7]. That is, like today's news broadcasts, simultaneous sign language interpretations are inserted, but this still does not address the issue of interaction for deaf people very well. This is because to this day, deaf education is still a phase of discovery and exploration, and there is no complete system worldwide, which results in the use and meaning of sign language not being identical in different countries and regions, and according to the World Hearing Impairment Alliance, there are currently more than 7 x 107 people worldwide who communicate in over 300 sign languages [1], just like the languages of different countries. A report drafted in France (Gillot, 1998) highlights that 80% of early deaf people are considered almost illiterate, even though they may have been exposed to literal vocal language. In fact, even if these deaf people succeed in mastering vocal language, processing vocal language content can be a rather tiring task for them.

Rastgoo (2021) argues that the best solution is to be able to transcribe the sign language in its entirety ('transcription' is the use of a graphic system to represent linguistic phenomena that occur in spoken or signed works), just like the International Phonetic Alphabet in vocal language, the sign The graphical system of sign language should capture each movement made by the signer in order to convey a meaning, and then enable an automatic translation(See Figure 3).[13]

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

Figure 3:American Sign Language Finger-Spelling of "ANXIETY"

The ultimate goal of most computer scientists currently working in the field of deaf interaction is to build an automatic interface that enables full translation between spoken and signed language, which, as Bragg (2019) puts it, is a mapping of 1D (verbal) streams to 4D (spatially and temporally symbolic) streams, and vice versa. The reality, however, is that automatic interpretation of sign language, at this stage, remains an illusion.[9]

4. How to improve the efficiency of sign language recognition

But returning to the technology, the approach to sign language recognition essentially relies on dataset-driven [2], and although relatively rich databases have been established in countries such as China, Germany, the USA and Iran [3], much of the research on sign language recognition technology at this stage is still stuck in the laboratory context of video understanding, and the research that can really extend sign language recognition technology to address recognition challenges in complex environments such as lighting, perspective changes, and hand occlusion in real time.

However, in the near future, the accuracy of recognition will no longer be limited by the complexity of the real world, and after 5-20 years of development there will inevitably be large-scale datasets of hearing-impaired people recorded in real environments. Secondly, the diversity of data formats, recording personnel, annotation features and scenarios can also greatly enhance the robustness of recognition methods, by which is meant the ability of a control system to maintain certain performance characteristics under certain (structure, size) parameter regimes [4]. This means that the user will not have to worry about accurate recognition due to the different expressions of natural sign language in different geographical areas or the complexity of the surrounding environment, as all sign language data will be recorded in a shared dataset.

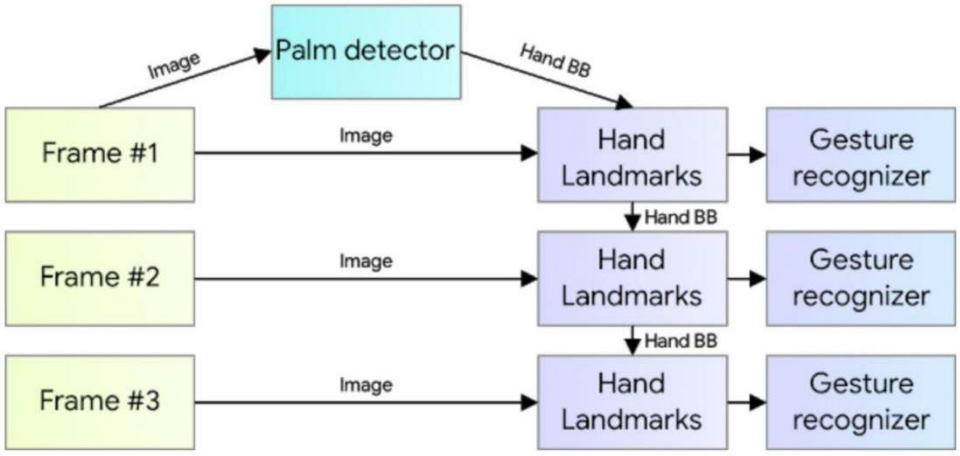

Figure 4:Vision-Based Sign-Language Recognition Pipeline: Input Frames → Palm Detection → Hand Landmark Extraction → Gesture Classification

Secondly, research on the extensibility model of sign language knowledge will also be gradually improved. The model can have a powerful ability to expand and train huge amounts of sign language data in real-life applications on a case-by-case basis, which Mujahid (2021) argues will lead to (See Figure 4).[5]

(a) Near-domain and cross-domain transfer learning methods extend the range of training data refinement to help alleviate sign language annotation bottlenecks.

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

(b) The recognition accuracy of zero-sample learning models is sufficient for applicability, and zero-sample learning research conducted in the area of continuous sign language recognition remains well established.

(c) Advances in lifelong learning mechanisms for sign language recognition continue to expand the knowledge base of models as they are deployed and applied.

5. Application in life scenarios

5.1. Applications in everyday life

But back in real life, developing software applications or solutions for users with disabilities can be a challenge, even for experienced user interface designers. Yet the general public is more interested in using sign language recognition on portable devices such as mobile phones, so developing practical models that combine lightweight, high-speed reasoning with multimodal multi-trait fusion for rapid deployment, real-time recognition and robustness is an indispensable project for the future.

Many people may think that writing directly or typing on a mobile phone is the answer to this problem, so why spend so much time and effort on sign language recognition when, for the vast majority of people, communication through speech is a matter of course. But for deaf people, sign language is the way they communicate with each other, so sign language recognition is necessary in all aspects of life. And sophisticated sign language recognition technology can solve a lot of problems for them. The connected device can be a watch, a finger holder or a ring and other simple gestures. For example, when they go to the hospital for a consultation they don't have to worry about the doctor not understanding the sign language, they just need to sign in front of the doctor and then the sign language recognition device can extract the gesture features and analyse them, then match them with the existing data set and finally communicate in an unobstructed way by converting them into voice announcements.

Since communication is no longer the biggest obstacle, there are more employment options for deaf people. The current employment situation for this group is bleak, with the majority being unemployed or non-working, while unaddressed hearing loss costs the world $980 billion annually.This includes health sector costs (excluding hearing device costs), educational support costs, lost productivity and social costs, 57% of which are attributable to low and middle income countries. Even when successful in employment, the jobs are relatively homogeneous and boring and poorly paid, such as carpentry, flower arranging and massage. Once applied, sign language recognition can be carried out to allow them to communicate in a simple way. Jobs such as shopping mall staff, cleaning staff and others that do not require complex communication can be filled. These special groups stand in the sunlight and confidently communicate with customers in sign language, while ordinary people can also convert the sounds they make into sign language through translation equipment so that they can understand what the other person is saying.

5.2. Applications in the field of education

However, the development of the sign language profession has been uneven around the world, with the exception of the first countries to develop sign language interpretation and education, such as the UK and Germany, mentioned above, where it is rare, and in developing countries, children with hearing loss and deafness often do not receive schooling. And with the exception of a few teacher training colleges and universities that offer sign language courses in special education, sign language education, especially social training, is weak and lacks clear, unambiguous and uniform criteria for access and evaluation.

In the future, the multimodal model of sign language will be used in basic education to facilitate the understanding and learning of sign language by deaf people. The development of this model will enable the linking of sign language movements with schematic pictures and text, allowing beginners to understand the semantics of sign language more deeply. And then, based on the sign language multimodal model, it can further combine education and testing to alleviate the shortage of sign language teachers to some extent. Teachers are allowed to create application tests to further put pressure on students to take the test. To facilitate learning, word predictions can be made when writing text and the ability to use audio and video clips as well as sign language displays for learning [10]. At the interactive level, technologies that enable automatic text-to-image generation, real-time sign language

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

vocabulary recognition, expression and lip recognition can also help deaf people to receive sign language education in a more comprehensive and systematic way. Future sign language recognition technologies can therefore be applied not only to special education schools, institutions offering teacher training in sign language teaching, but also to places such as general sign language teaching sites and grassroots training sites of the Disabled Persons' Federation. This environment and the tools presented in this paper create an ideal opportunity to truly understand the needs of people with disabilities, especially children, and the adaptation of these tools to their needs, and to continue to experiment with new solutions or software for children with disabilities in the future.

6. Sign Language Recognition and AI

In addition to machine learning-based methods for sign language recognition, AI (deep learning) is another approach, but research is significantly more difficult than machine learning as the technology is not yet used on a large scale. And looking back at the modern state of life, one finds that the handling of public affairs is now increasingly being replaced by voice interaction, image recognition and other AI technologies, while hearing and visually impaired groups have more trouble accessing services. at the end of 2018, the UN published a report on the UK government's digital outcomes, resulting in data showing that in England, homelessness has increased by 60 %, that there are 1.2 million people on waiting lists for secure housing, and that the demand for food banks used to provide relief to the poor has nearly quadrupled - because many people in poverty don't know how to apply for poverty benefits on the internet, or even have an internet connection at home, and end up sinking deeper and deeper into poverty. So while AI is facilitating human life, it should also take into account special groups, and these conveniences are perhaps making them more and more distant from ordinary life, so applying AI technology to sign language recognition is a trend that makes technology less cold and more humane.

With the current technology, the initial application of AI's gesture recognition capability has already begun in some software, such as "comparing hearts" when taking photos to trigger some AR effects. By capturing these gestures and matching them with the semantic meaning of the gestures, it will be possible to translate and generate sign language in the future. The sign language AI technology itself can be divided into two solution paths: recognition models and datasets. In terms of datasets, the future can build its own sign language recognition dataset through contact with socially relevant organisations and hearing impaired people, and this dataset can be expanded for local differences in sign language expressions, in terms of expression habits and speed.

As for recognition models, scientists have also proposed newer algorithm building concepts, such as extracting static and dynamic information from gestures through 2D and 3D convolutional neural networks, respectively, to enhance video recognition through integrated processing, freeing it from the shackles of other sensors. Also for the phenomenon of whole sentence expressions in sign language, word-level information mining is added at the end of the video frame to verify the information presented by the feature extractor and further go on to determine the gesture-to-word expression boundaries, which, in addition to improving recognition accuracy, also improves the ability to summarise regional expressions in natural sign language. On this basis, contextual understanding can also be introduced into the algorithm model in order to face more complex sign language recognition translation needs.

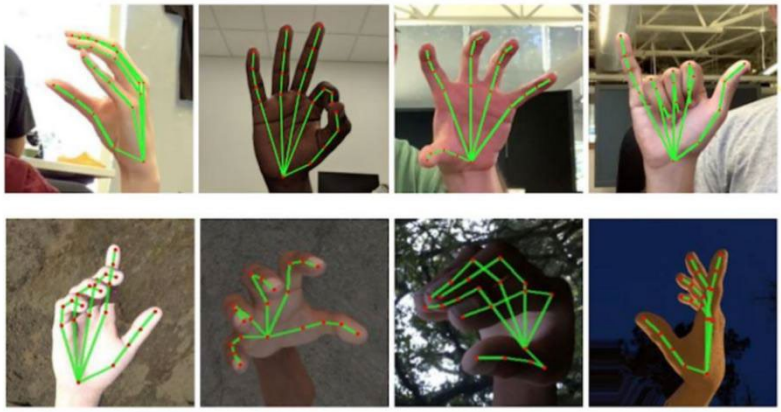

Figure 5:Examples of Hand Landmark Detection in Diverse Lighting and Background Conditions

At the same time the technology, as it can be improved, will not be limited at the application

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

scenario end in the future. Because of the high computing power requirements of high-precision algorithms, AI sign language translators currently need to rely on the background computing of high-performance computers; and because recognition is done through image video, the recognition of complex scenes is not yet high. Therefore, future AI sign language interpreters can gradually improve the recognition of various complex scenes, whether it is a quiet home or a crowded public place, such as the airport, high-speed railway, civil affairs, etc., where the hearing-impaired can communicate without barriers, to help build a city with barrier-free information with technology(See Figure 5).

7. Summary

The problems faced by deaf people in some developing countries are more severe than those in developed countries. According to the United Nations Health Organisation, nearly 80 per cent of people with disabling hearing loss live in low- and middle-income countries. Social welfare protection is low, and there is little scope for employment and job opportunities, as well as little access to education. Although the government is continuing to promote the development of programs to assist the elderly and disabled, there are few devices and applications on the market that are truly adapted for people with hearing loss. Although in some cases the applications that are not yet fully developed do not meet the needs of the average user in terms of natural interaction, for special groups there is an urgent need for these technologies to serve them, as natural human-computer interaction can make a difference to the nature of their agency, allowing them to use computers to communicate with the outside world, to learn and improve or to relax, providing them with access to the digital world It also allows digital resources to be truly accessible to all. Every advancement in these areas of application can improve the human-computer interaction experience for special groups to some extent, so attempts in these areas are important and valuable.

The current assisted sign language to speech tool for the hearing impaired is based on the iOS platform and is mainly based on serving the hearing impaired to achieve normal emotional communication, providing them with instant camera, sign language to emotional speech service. As the internet and information technology progresses towards the goal of broadband, mobile devices are becoming more popular and are gradually becoming more and more involved in all aspects of people's daily lives. Although the relatively low purchasing power of special groups and the different needs of different situations lead to high development costs for the corresponding human-computer interaction support systems for special groups. But I believe that in the future, mature sign language recognition technology can not only enter the daily life of deaf people and make it affordable for everyone, but also allow them to express their emotions more directly, after all, solving the communication is only the first step to let them experience a wonderful life, the future really let them and ordinary people, useful and beautiful life and not be looked at differently is a long way to go.

References

*[1] Murray, J. (2020). *World Federation of the Deaf*. http://wfdeaf.org/our-work/*

*[2] Tao, T., & Liu, T. (2022, December 14). A review of sign language recognition techniques based on sign language expression content and expression features. *Journal of Electronics and Information*, 1–19. http://kns.cnki.net/kcms/detail/11.4494.TN.20221107.1650.015.html [In Chinese]*

*[3] De Amorim, C. C., & Zanchettin, C. (2022). ASL Skeleton3D and ASL-phono: Two novel datasets for the American Sign Language. *arXiv preprint arXiv:2201.02065*.*

*[4] Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2), 144–152.*

*[5] Mujahid, A., Awan, M. J., Yasin, A., Mohammed, M. A., Damaševičius, R., Maskeliūnas, R., & Abdulkareem, K. H. (2021). Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences*, 11(9), 4164. https://doi.org/10.3390/app11094164*

*[6] Rømen, D., & Svanæs, D. (2012). Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society*, 11(4), 375–385.*

*[7] Brajnik, G., Yesilada, Y., & Harper, S. (2010, October). Testability and validity of WCAG 2.0: The expertise effect. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 43–50). https://doi.org/10.1145/1878803.1878813*

*[8] Kalyanathaya, K. P., Akila, D., & Rajesh, P. (2019). Advances in natural language processing – a survey of current research trends, development tools and industry applications. *International Journal of Recent Technology and Engineering*, 7(5C), 199–202.*

ISSN 2618-1568 Vol. 7, Issue 4: 75-82, DOI: 10.25236/FAR.2025.070411

*[9] Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., ... & Ringel Morris, M. (2019, October). Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 16–31). https://doi.org/10.1145/3308561.3353774*

*[10] Sobota, B., Korečko, Š., Hudák, M., Ondra, F., & Sivý, M. (2020, November). Disabled people and pictogram-text application with gesture recognition extension. In *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)* (pp. 640–645). IEEE. https://doi.org/10.1109/ICETA51985.2020.9379262*

*[11] Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language*. De Gruyter Mouton.*

*[12] Kyle, J. G., Kyle, J., Woll, B., Pullen, G., & Maddix, F. (1988). *Sign language: The study of deaf people and their language*. Cambridge University Press.*

*[13] Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign language recognition: A deep survey. *Expert Systems with Applications*, 164, 113794. https://doi.org/10.1016/j.eswa.2020.113794*