

Sign Language Recognition Techniques

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Bader Alsharif , Ali Salem Altaher , Ahmed Altaher , Mohammad Ilyas , Easa Alalwany

Historically, individuals with hearing impairments have faced neglect, lacking the necessary tools to facilitate effective communication. Building a sign language recognition system using deep learning technology plays a vital role in interpreting sign language to ordinary individuals and the reverse. This system would ease the process of communication between deaf and normal people.

image-based

American sign language

deep learning

transfer learning

1. Introduction

Throughout history, humans have employed a variety of communication techniques including gesturing, sounds, drawing, writing, and speaking. However, for people with deafness or hearing impairments, sign language is the primary means of communication and interaction with others, breaking down all barriers of hearing loss condition which severely limits their verbal communication. Because of communication barriers, people with these disabilities have fewer opportunities for development. Sign language is a spontaneous non-verbal language expressed by using manual gestures, facial expressions, and body language to convey messages and meaning. These signs may vary from one country to another, although they have some similarities in using sign language. Unfortunately, there is no universal sign language that can be used for all people with hearing impairments around the world ^[1]. According to the World Health Organization's (WHO) most recent research, 5% of the world's population—432 million adults and 34 million children—have disabling hearing loss, not to mention more than 1 billion people are susceptible to hearing loss due to extended and excessive to loud sounds ^[2]. In fact, people who lose their hearing sense under any circumstances will lose their ability to speak. The enormous number of deafness and hearing loss conditions has garnered the attention of many researchers and developers in the field of speech recognition and other multidisciplinary fields to conduct their study to assist people with hearing impairments. Their goal is to facilitate the daily life of people with hearing disabilities for communication and social interaction with other individuals. Consequently, with the rapidly growing deaf community, building a sign language recognition system using deep learning technology plays a vital role in interpreting sign language to ordinary individuals and the reverse. This system would ease the process of communication between deaf and normal people. As a result, people with hearing impairments will have the opportunity to become more engaged in society, developing social interaction and relationships ^{[3][4]}. Nowadays, the sole means for people with hearing impairments to communicate with other ordinary people is through interpreters. However, it is very costly to hire interpreters who have expertise in interpreting for the deaf, because of the limited number of such interpreters. Nevertheless, there are several obstacles in implementing a sign language recognition system to support the deaf and hearing loss community that should be discussed. Firstly, not all hearing-impaired individuals use sign language as a method of communication,

which may give them a sense of isolation and depression. Secondly, there are more than 200 different sign languages and dialects from different countries which may delay the process of implementing a sign language recognition system which would be applicable in various countries [4][5]. Lastly, not everyone is proficient in using today's modern technology due to a lack of education and development, which can be neglected. Numerous studies and research should be oriented to address and comprehend the obstacles that deaf and hard-of-hearing people encounter, which hinder their societal engagement [6].

2. Sign Language Recognition Techniques

This section explores an influx of related publications on sign language recognition techniques. According to [7][8][9][10], the implementation of a sign language recognition system can be carried out either by using a sensor-based approach, an image-based approach, or both approaches (hybrid), as can be seen in **Figure 1**.

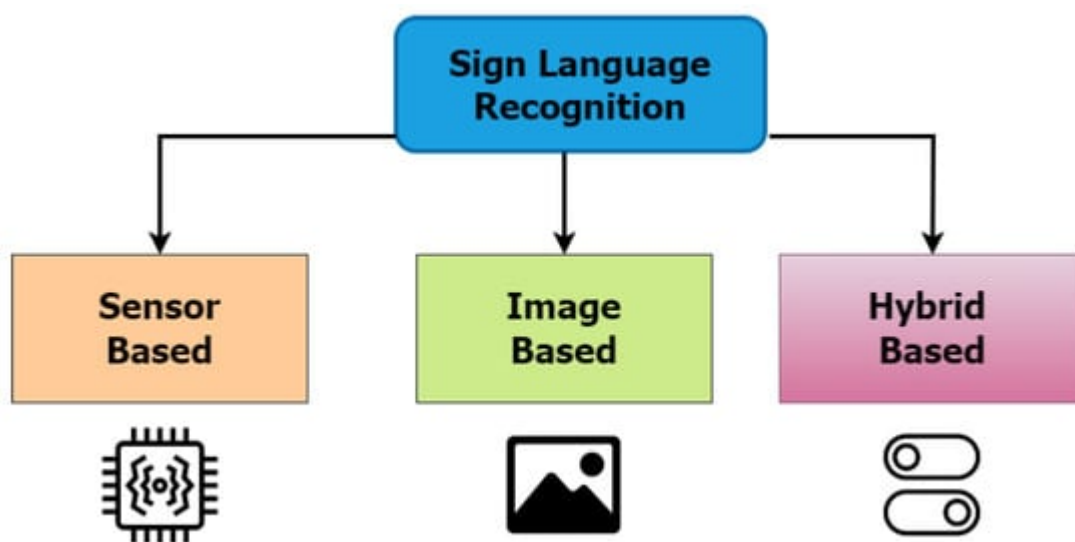


Figure 1. Main approaches for sign language recognition system.

In the sensor-based system technique, the user wears a specialized glove equipped with multiple sensors and wires. These sensors assist the system in tracking and recording the movements of hands and fingers. The information transmitted to a computer includes data on finger bending, movements, orientation, rotation, and hand position for interpretation. This interaction between the smart glove and computer is a clear example of human–computer interaction. There are two types of sensor-based methods used in sign language acquisition: sensors that can only detect finger bending and sensors that detect hand motion and orientation [10][11]. For more detailed information, readers can refer to the comprehensive survey paper “Systems-based sensory gloves for sign language recognition” [10].

In an image-based approach, there is no need to wear a glove overloaded with wires, sensors, and other materials. The idea behind image-based systems is to use image processing techniques and algorithms to perceive sign gestures [12][13][14]. Image-based sign language recognition systems can be developed using smart devices since

most smart devices have high-resolution cameras that allow natural movements and easy availability. Sensor-based systems are accurate and reliable because they simulate hand gestures. Nevertheless, sensor-based techniques have significant drawbacks, such as the user's heavy glove size making it uncomfortable to wear [7][10][15]. In addition, the glove has several wires connected to a computer, which limits the user's mobility and its usage of real-time applications [10].

In [16], the authors developed an Arabic sign language (ArSL) recognition system based on a CNN. A CNN is a sort of artificial neural network (ANN) used in deep learning for image processing, recognition, and classification. The system's implementation recognizes and translates hand gestures into text to bridge the communication gap between deaf and non-deaf people. They used a dataset consisting of 40 Arabic signs, with each sign having 700 different images, which is a principal factor for training systems to have multiple samples per sign. They employed various hand sizes, lighting, skin tones, and backgrounds to increase the system's dependability. The result showed an accuracy of 97.69% for training data and 99.47% for testing data. The system was successfully implemented in both mobile and desktop applications. In the same context, Ref. [17] introduced an offline ArSL recognition system based on a deep convolutional neural networks model that can automatically recognize letters and numbers from one to ten. They utilized a real dataset composed of 7869 RGB images. The proposed system achieved an accuracy of 90.02% by training 80% of dataset images. The research introduced in [18] aims to translate the hand gestures of two-dimensional images into text using a faster region-based convolutional neural network (R-CNN). Their system mapped the position of the hand gestures and recognized the letters. They used a dataset of more than 15,360 images with divergent backgrounds that were captured using a phone camera. The result shows a recognition rate of 93% for the collected ArSL images dataset. The goal of this proposed study by [19] is to create a system that can translate static sign gestures into words. They utilized a vision-based method to obtain data from a 1080 full-HD web camera of the signer. The camera will capture only the hands to feed into the system. The dataset will be built through continuous capturing. CNN is applied as a recognition method for their system. After training the model and testing it, the system acquired an average of 90.04% accuracy for recognizing the American Sign Language (ASL) alphabet, 93.44% for numbers (from 1 to 10), and 97.52% for static word recognition. In [20], the authors presented a vision-based gesture recognition system that uses complicated backgrounds. They designed a method for adapting to the skin color of different users and lighting conditions. Three types of features were combined: principal component analysis (PCA), linear discriminant analysis (LDA), and support vector machine (SVM) to describe the hand gestures. The dataset used contains 7800 images for the ASL alphabet. The overall accuracy achieved is 94%. The authors in this work [21] utilized a supervised ML technique to recognize hand-gesturing in ArSL using two sensors: Microsoft's Kinect with a Leap Motion Controller in a real-time manner. The proposed system matched 224 cases of the Arabic alphabet letter signed by four participants, each of whom performed over 56 gestures. The work carried out by [22] presents a visual sign language recognition system that automatically converts solitary Arabic word signs into text. The proposed system has four main stages: hand segmentation, hand tracking, hand feature extraction, and hand classification. The use of the hand segmentation technique is performed to utilize dynamic skin detectors. Then, the segmented skin blobs are used to track and identify the hands. This proposal uses a dataset of 30 isolated words frequently used by hearing-impaired students daily in school. The result shows that the system has a recognition rate of 97%. In [23],

the authors created a dataset and a CNN sign language recognition system to interpret the American sign gesture alphabet and translate it to our natural language. Three datasets were used to compare the results and accuracy of each. The first dataset, which belongs to the authors, has 104,000 images for 26 letters of ASL; the second dataset of ASL has 52,000 images; and the third dataset contains 62,400 images. The datasets were split into 70% for the training sets and 15% each for the validation and testing sets. The overall accuracy for all three datasets based on the CNN model is 99% with a slight difference in the decimal values. For another proposed sign language recognition system, ref. [15] trained a CNN deep learning model to recognize 87,000 ASL images and translate them into text. They were able to achieve an accuracy of 78.50%. For another ASL classification task, ref. [24] developed EfficientNet model to recognize ASL alphabet hand gestures. Their dataset size was 5400 images. They achieved an accuracy of 94.30%. In the same context, [25] used the same dataset of 87,000 images for classification. They used (AlexNet and GooLeNet) models, and their overall training results were 99.39% for AlexNet and 95.52% for GoogLeNet. In [26], the authors evaluated ASL alphabet recognition utilizing two different neural network architectures, AlexNet and ResNet-50. The results showed that AlexNet achieved an accuracy of 94.74%, while ResNet-50 outperformed it significantly with an accuracy of 98.88%.

References

1. Zwitserlood, I.; Verlinden, M.; Ros, J.; Van Der Schoot, S.; Netherlands, T. Synthetic signing for the deaf: Esign. In Proceedings of the Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment, CVHI, Granada, Spain, 29 June–2 July 2004.
2. World Health Organisation. Deafness and Hearing Loss. 2021. Available online: <http://https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> (accessed on 8 November 2022).
3. Alsaadi, Z.; Alshamani, E.; Alrehaili, M.; Alrashdi, A.A.D.; Albelwi, S.; Elfaki, A.O. A real time Arabic sign language alphabets (ArSLA) recognition model using deep learning architecture. *Computers* 2022, 11, 78.
4. Alsharif, B.; Ilyas, M. Internet of Things Technologies in Healthcare for People with Hearing Impairments. In Proceedings of the IoT and Big Data Technologies for Health Care: Third EAI International Conference, IoTCare 2022, Virtual Event, 12–13 December 2022; Proceedings. Springer: Berlin/Heidelberg, Germany, 2023; pp. 299–308.
5. Farooq, U.; Rahim, M.S.M.; Sabir, N.; Hussain, A.; Abid, A. Advances in machine translation for sign language: Approaches, limitations, and challenges. *Neural Comput. Appl.* 2021, 33, 14357–14399.
6. Latif, G.; Mohammad, N.; AlKhalaf, R.; AlKhalaf, R.; Alghazo, J.; Khan, M. An automatic Arabic sign language recognition system based on deep CNN: An assistive system for the deaf and hard of hearing. *Int. J. Comput. Digit. Syst.* 2020, 9, 715–724.

7. Mummadi, C.K.; Philips Peter Leo, F.; Deep Verma, K.; Kasireddy, S.; Scholl, P.M.; Kempfle, J.; Van Laerhoven, K. Real-time and embedded detection of hand gestures with an IMU-based glove. *Informatics* 2018, 5, 28.
8. Elons, A.; Ahmed, M.; Shedid, H.; Tolba, M. Arabic sign language recognition using leap motion sensor. In *Proceedings of the 2014 9th International Conference on Computer Engineering & Systems (ICCES)*, Cairo, Egypt, 21–23 December 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 368–373.
9. Kammoun, S.; Darwish, D.; Althubeany, H.; Alfull, R. ArSign: Toward a mobile based Arabic sign language translator using LMC. In *Proceedings of the Universal Access in Human–Computer Interaction. Applications and Practice: 14th International Conference, UAHCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, 19–24 July 2020; Proceedings, Part II 22; Springer: Berlin/Heidelberg, Germany, 2020; pp. 92–101.*
10. Ahmed, M.A.; Zaidan, B.B.; Zaidan, A.A.; Salih, M.M.; Lakulu, M.M.B. A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. *Sensors* 2018, 18, 2208.
11. El-Alfy, E.S.M.; Lugman, H. A comprehensive survey and taxonomy of sign language research. *Eng. Appl. Artif. Intell.* 2022, 114, 105198.
12. Alzohairi, R.; Alghonaim, R.; Alshehri, W.; Aloqeely, S. Image based Arabic sign language recognition system. *Int. J. Adv. Comput. Sci. Appl.* 2018, 9, 3.
13. Al Ani, L.A.; Al Tahir, H.S. Classification Performance of TM Satellite Images. *Al-Nahrain J. Sci.* 2020, 23, 62–68.
14. Nazari, K.; Ebadi, M.J.; Berahmand, K. Diagnosis of alternaria disease and leafminer pest on tomato leaves using image processing techniques. *J. Sci. Food Agric.* 2022, 102, 6907–6920.
15. Mohandes, M.; Liu, J.; Deriche, M. A survey of image-based arabic sign language recognition. In *Proceedings of the 2014 IEEE 11th International Multi-Conference on Systems, Signals & Devices (SSD14)*, Barcelona, Spain, 11–14 February 2014; pp. 1–4.
16. Al-Obodi, A.H.; Al-Hanine, A.M.; Al-Harbi, K.N.; Al-Dawas, M.S.; Al-Shargabi, A.A. A Saudi Sign Language recognition system based on convolutional neural networks. *Build. Serv. Eng. Res. Technol.* 2020, 13, 3328–3334.
17. Hayani, S.; Benaddy, M.; El Meslouhi, O.; Kardouchi, M. Arab sign language recognition with convolutional neural networks. In *Proceedings of the 2019 International Conference of Computer Science and Renewable Energies (ICCSRE)*, Agadir, Morocco, 22–24 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.
18. Alawwad, R.A.; Bchir, O.; Ismail, M.M.B. Arabic sign language recognition using Faster R-CNN. *Int. J. Adv. Comput. Sci. Appl.* 2021, 12, 3.

19. Vanaja, S.; Preetha, R.; Sudha, S. Hand Gesture Recognition for Deaf and Dumb Using CNN Technique. In Proceedings of the 2021 6th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 8–10 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–4.
20. Pan, T.Y.; Lo, L.Y.; Yeh, C.W.; Li, J.W.; Liu, H.T.; Hu, M.C. Real-time sign language recognition in complex background scene based on a hierarchical clustering classification method. In Proceedings of the 2016 IEEE Second International Conference on Multimedia Big Data (BigMM), Taipei, Taiwan, 20–22 April 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 64–67.
21. Almasre, M.A.; Al-Nuaim, H. A real-time letter recognition model for Arabic sign language using kinect and leap motion controller v2. *Int. J. Adv. Eng. Manag. Sci.* 2016, 2, 239469.
22. Ibrahim, N.B.; Selim, M.M.; Zayed, H.H. An automatic Arabic sign language recognition system (ArSLRS). *J. King Saud Univ.-Comput. Inf. Sci.* 2018, 30, 470–477.
23. Kasapbaşı, A.; ELBUSHRA, A.E.A.; Omar, A.H.; Yilmaz, A. DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals. *Comput. Methods Programs Biomed. Update* 2022, 2, 100048.
24. AlKhuraym, B.Y.; Ismail, M.M.B.; Bchir, O. Arabic sign language recognition using lightweight cnn-based architecture. *Int. J. Adv. Comput. Sci. Appl.* 2022, 13, 4.
25. Cayamcela, M.E.M.; Lim, W. Fine-tuning a pre-trained convolutional neural network model to translate American sign language in real-time. In Proceedings of the 2019 International Conference on Computing, Networking and Communications (ICNC), Honolulu, HI, USA, 18–21 February 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 100–104.
26. Ma, Y.; Xu, T.; Kim, K. Two-Stream Mixed Convolutional Neural Network for American Sign Language Recognition. *Sensors* 2022, 22, 5959.

Retrieved from <https://encyclopedia.pub/entry/history/show/113163>