Derleme Makalesi
Bayburt Üniversitesi Fen Bilimleri Dergisi
Cilt 4, Sayı 1

Geliş Tarihi: 12.02.2021
Kabul Tarihi: 26.04.2021

dergipark.gov.tr/bufbd

Arabic Sign Language Recognition: A Review

Ghadir Abdulhakim Abdo Abdullah ALSELWI*, Tuğrul TAŞCI

Sakarya University, Information Systems Engineering Department, Sakarya, Turkey.

Keywords:

Sign Language (SL), Arabic Sign Language (ArSL), Arabic Sign Language Recognition (ArSLR)

Abstract

Sign language is the language used by deaf people to communicate with one another and with common people. An interpreter is typically required when an ordinary person wishes to communicate with a deaf person. Visual sign language recognition is a dynamic area of research in computer vision. A functional sign language recognition (SLR) system can provide an opportunity for a deaf person to communicate with non-signed persons without the need for an interpreter. It can be used to produce speech or text that makes the mute more autonomous. Arabic Sign Language Recognition (ArSLR) is an important area of study. It helps to eliminate barriers between the deaf and the community. This paper includes a review of the Arabic Sign Language Recognition System (ArSLRS). Present challenges and future research opportunities are also illustrated.

1 INTRODUCTION

There are more than 300 sign languages around the world that 70 million deaf people use according to the World Federation of the Deaf [21]. Many people with impairments in speech and hearing can't read or write in their normal languages. Sign language (SL) is the mother tongue used in contact with deaf people. SL mostly relies on gestures to convey meaning [28]. SL consists of the use of multiple body components such as fingertips, wrist, arm, head, body, and facial expression [9]. SL, as a systemic hand gesture of facial gestures and signs, is therefore used as a communication device to support the everyday contact between the surrounding and speech-impaired community.

Hearing impairment is the concept of the board referred to as partial or total hearing loss in one or both ears. The degree of disability ranges from mild, moderate, profound, or severe [14]. People with hearing or speech impairments are dependent mainly on SL as a means of daily contact [29]. SL is the most formal in contrast with other gestures. It has a wide variety of signs and a particular meaning for each symbol. A functional sign language recognition (SLR) system can provide an opportunity for a deaf person to communicate with non-signed persons without the need for an interpreter.

The organization of the paper is: Section 2 gives information about SLs and SL in Arabic countries. Section 3 focuses on a typical SL recognition system. Section 4 describes the typical vision-based hand gesture recognition system with various methods/ techniques available in the literature. Section 5 discusses the challenges of developing the ArSLR system. Section 6 contains a discussion and conclusion.

2 ARABIC SIGN LANGUAGE (ARSL)

Each nation or area currently has its SL. Therefore In terms of syntax and grammar, these SLs vary from country to another. For instance, Australian SL is known as Australia's Sign Language (Auslan) and Japanese Sign Language (JSL), whereas Arabic Sign Language (ArSL) is used in the Arab world [27, 12, 33]. The Arabic language is very important; thus it is the fourth spoken language in the world [18].

ArSL is the native language commonly practiced by the deaf community of Arabic communities. Unlike spoken languages, ArSL allows a deaf individual to convey meanings and ideas by gestures instead of sounds. In 2001 ArSL was officially launched by the Arab Federation of the Deaf. However, ArSL is still in its growth process, the Arabic language being one of the main languages around the world. "Diglossia" is perhaps the biggest

*e-Posta: abdulhakim.alselwi@ogr.sakarya.edu.tr

challenge faced by ArSL. In reality, in every country, instead of writing language, regional dialect is spoken. In this way, various spoken dialects made various ArSLs. They are as many as Arab nations; however, there are numerous words in common and similar alphabet [22].

In recent years ArSLs are in their developmental phases. Besides, people have become aware of the presence of groups with hearing disabilities. There are almost closed Arab Deaf communities due to a little contact between a deaf and a hearing community and this is focused primarily on families, relatives, friends, and professionals of deaf members. ArSL relies on the shapes of the letters. It, therefore, contains letters which are distinct from the different languages of the representation of the alphabet [1].

Arabic is a complex and comfortable language, but at the same time, it is one of the Semitic dialects spoken by almost 380 million people across the globe as their original official language [23], [24]. Arabic is the official language of the regions from North Africa to the Arabian Gulf. The Arabic language is facing a variety of difficulties. As shown in Figure 1. ArSL is made up of 28 different characters, formed from left to right. Various types of techniques and classifiers are used for the ArSL recognition system (ArSLRS) [1]. Generally, the SL recognition involves several phases of the process, such as extraction of features, segmentation, and classification [22].

Figure 1. ArSL Alphabets

3 ARSLRS ARCHITECTURE

SL Recognition System (SLRS) is one of the application areas of human-computer interaction (HCI) where signs of hearing-impaired people are converted to text or voice of the oral language [14]. ArSL recognition research efforts can be split into two approaches, such as Vision-Based Recognition (VBR) and Sensor-Based Recognition(SBR). The Arabic alphabet, isolated words, and continuous SL recognition activities are usually applied to these two systems. The VBR is the most common approach in ArSLRS compared to the SBR approach [6]. Therefore, we will only review the publications that used VBR.

The recognition of SL is very significant not only from an engineering point of view but also in terms of its effect on humans and society. SLRS operates on five main parameters: hand shapes; hand movement; hand and head orientation; hand and head position; and facial expressions [25].

The structure of the recognition process is shown in Figure 2. First of all, a gesture picture is created using a camera that is connected to the device or by dataset preprocessing the acquired signs to make them convenient for the next levels. Each sensing technology differs across several dimensions, including precision, resolution, latency, range of motion, user comfort, and cost. The tracker often has to deal with shifting shapes and sizes of a gesture-generating object (which varies between individuals), other moving objects in the background, and noise [25].

Then the image enters the stage of segmentation which is crucial because all the following steps rely on a correct segmentation of the image. By segmenting the image, the system can use the region corresponding to compute properties necessary for the extraction of the image's features. Using the results of the image segmentation stage, the extraction stage of the feature transforms the image into a collection of features needed for the process of recognition which is performed in the last stage. Selecting good features is a vital step in every object recognition system. The purpose of the feature extraction process is to represent the image by a collection of numerical features that correspond to useful details and to eliminate the redundancy of the image data. In the last phase, the extracted features from the image data are recognized as a specific gesture [3].

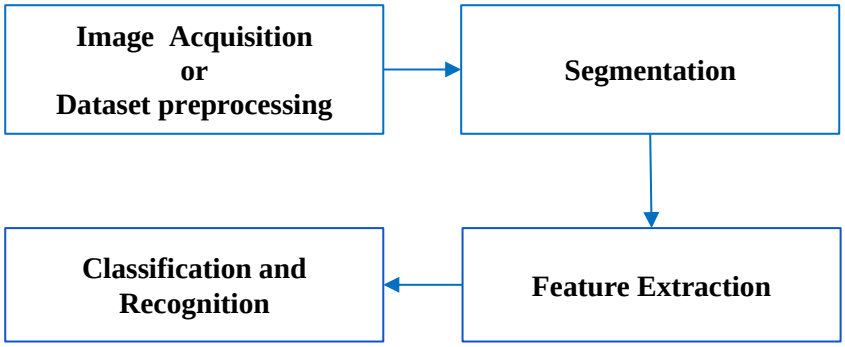

Figure 2. Basic block diagram of SLRS.

4 RELATED WORK

A variety of published scientific works have been published in different years, according to the research work on ArSLRS. In 2001 Al-Jarrah and Halawani developed a system for automatic translation of gestures of the manual alphabets in ArSL by building a series of Adaptive Neuro-Fuzzy Inference System (ANFIS) networks. This system was able to recognize the Arabic alphabets with a 93.55% recognition rate [3].

In 2009, Al-Rousan demonstrated a way to recognize ArSL using a single video camera and introduced the first automatic ArSLRS based on Hidden Markov models (HMMs). In signer-dependent instances, the device obtains a word recognition rate of 98.13%, 96.74%, and 93.8%, on rain data in offline mode, on test data in offline mode, and test data in online mode, respectively. On the other hand, the device obtains a word recognition score of 94.2% for signer-independent cases and 90.6% for offline and online modes respectively [5].

In 2010 Shanableh and Assaleh introduced two novel approaches to the extraction of features applied to the video-based ArSLR, namely motion representation through motion estimation and motion residuals. Comparisons with current work indicate that up to 39% of misclassifications have been corrected [8].

Also, in 2010, Tolba proposed a new technique and addressed the problems of the standard generation of image signatures and their standardization for the recognition of images and static alphabets using Pulse Coupled Neural Network (PCNN) [30]. In 2013 Tolba also focused on how to identify, segment, and categorize the connected sequence of gestures in real-time using the PCNN and graph-matching method with a 70% recognition rate [31]. Moreover, in 2013 Elons proposed a novel technique for dealing with pose variations in 3D object recognition by using PCNN to generate image features from two different viewing angles. They proposed a technique that uses two 2D image features to construct optimized 3D quality features with a 90% recognition rate. Elons then proposed a new approach to increase the quality of the features by using PCNN, Continuity Factor which got a 90% recognition rate [10].

Previous attempts at automated vision-based recognition of ArSL concentrated primarily on finger spelling and recognition of isolated gestures. In 2010, the first continuous ArSL was documented by building on established research into the extraction of features and pattern recognition. This system is based on spatial-temporal extraction features and HMMs with a 94% recognition rate [8].

In 2011 by using user-independent classification (DCT, KNN, and polynomial). Shanableh and Assaleh filtered out user-dependent information and presented a solution for their recognition of isolated ArSL with an 87% classification rate [27].

In 2012 Mohandes used face detection, geometric features, and HMM to propose an image-based system that has been trained to work in both signer-dependent and signer-independent modes with a 95% recognition rate [20].

Most of the methods studied in ArSLR used machine learning and computer vision techniques for sign recognition. However, different frequency domain transformations (Viz. Fourier, Hartley, and Log-Gabor transforms) were evaluated for ArSLR by Luqman and Mahmoud. The results show that the Hartley transform is effective for the recognition of ArSLR with an accuracy of 98.8% using the SVM classifier [17].

In 2018, by using a dynamic skin detector and skin-blob tracking, Ibrahim presented an automated visual SLRS that translates isolated Arabic words into text with a 97% recognition rate [14]. Alzohairi also automatically recognized ArSL alphabets using an image-based methodology with recognizing 63.5% of Arabic Alphabet gestures [7].

In 2020, Kamruzzan proposed a vision-based method by applying Convolutional Neural Network (CNN) for the identification of Arabic hand sign-based letters and translating them into Arabic speech with a 90% recognition rate [15]. A 97.6% recognition was obtained when deep CNN architectures have been trained and validated to classify signs of the Arabic alphabet [16]. After that Saleh and Issa utilized transfer learning and fine-tuning deep CNNs to boost the accuracy of recognizing. The resulting model has a validation accuracy of 99% [26]. Hisham and Hamoud introduced an ArSLR system that uses a Leap Motion Controller and Latte Panda. The suggested system is extended to 30 hand gestures consisting of 20 single-hand gestures and 10 double-hand gestures. Results of accuracy are AdaBoost (92.3% for single-hand, 93% for double-hand gestures) [13]. Alnahhas presented a creative method to recognize words in ArSL using a Leap Motion system that helps create a 3D human hand model using infrared and obtains an 89% classification rate for one-hand, 96% for two hands gestures [4]. Mohamed developed a computerized system which focused on depth-measuring cameras and computer vision techniques for capturing and segmenting images of facial expressions and hand gesture with a 90% recognition rate [19].

5 CHALLENGES IN SLRS

SL functions and spoken language functions are entirely different. SL is essentially based on the properties of space and iconicity. Hand parameters such as shape, movement, orientation, and position as well as facial expression, movement of the mouth are considered to understand the sign. These parameters take place concurrently and are articulated in space. Building a syntactic and semantic rule-based structure is necessary since one sentence in a spoken language can be expressed by a single-sign in SL [25].

Although automatic speech recognition has now progressed to the point that it is commercially available, automatic SLR and in particular, ArSL is still in its infancy. Automatic gesture recognition is a challenge, too[5]. Besides, publicly accessible databases are limited in both quality and quantity, making many conventional pattern recognition learning algorithms ineffective for building classifiers. However, the database is limited to ArSL alphabets and the use of other ArSL signs that require more motion makes the recognition task more difficult [17]. Building an automatic hand gesture recognition system in Arabic, solving the issue of recognition, and improving the method of real-time recognition is a challenge [13]. However, the primary challenge faced by any sign language recognition system is the ability to track the signer in the video of the signer with a variety of background clutter and different lighting conditions [25]. Researchers also reported some difficulties in terms of ArSL:

1. [13]
    a. A shortage in Arabic sign linguists
        i. No available document to build the corpus used for translation
            1. Corpus must be built from scratch.
    b. ArSL is very versatile and not universal; it differs according to different cultures and dialects.
    c. Very limited types of research on ArSL.
2. [16]
    a. Occlusions problem while performing sign.
    b. Adding more images to the dataset, since the number of participants who perform signs is limited.
3. [20]
    a. Camera quality, noise, illumination challenges.
    b. Lack of a standard or benchmark datasets for ArSL
        i. Makes it difficult to compare the findings with other works.
4. [14]
    a. No common databases available to researchers in the ArSLR field
        i. There will be a need to establish a database.
5. [7]
    a. Difficulty to recognize and segregate ArSL alphabet (similar gesture).
6. [10]

   a. Variations in geometry, photometry and viewing angle, noise, occlusions, and incomplete data.
   b. Some SLR systems lack extendibility and suffer from vulnerability to environmental factors, such as signer location and view angle.
7. [27]
   a. Co-articulation problem (the link between preceding and subsequent sign).
8. [10, 30].
   a. Identifying some postures

The computed-based recognition process for ArSL grammatical structures and the verification of sentences and their grammatical structures is difficult. ArSL has several gestures that are very similar to each other. Some important linguistic characteristics should be studied before the design of a rule-based framework (language processing engine).

6 DISCUSSION AND CONCLUSION

According to the research work on ArSLRS, a number of published scientific works have been published in various years. In our paper, we addressed some studies that were published between 2001 and 2020. Many approaches were used in these studies, including various neural network models, machine learning methods, classification algorithms, frequency domain transforms, and so on. In recent years, there has been a growing interest in the use of Leap Motion and neural network models such as CNN to achieve higher accuracy rates in ArSLRSs.

The main objective of this paper is to give the significance of ArSL as an interpretive language and to research the growth of ArSLRS across Arabic countries. Researchers are faced with a big problem with the availability of standard and common datasets. Most researchers are establishing and working on their own data sets. The deaf assistant system can bridge the communication gap between hearing-impaired and normal people without isolating them in society. It would be a great contribution to the Arab hearing-impaired by focusing on ArSLRSs so that they can become self-independent people, considering their deafness and muteness, they also can play a useful role in society.

Future research on ArSLRS may be conducted to develop the methodology to recognize overlapped gestures and build complete sentences out of SL gestures [4]. Hand tracking and segmentation can be enhanced to overcome the occlusion of the hand [19]. The most suitable descriptor for the recognition system remains an open research challenge [7]. Researchers may also examine how HMM can be used to identify a series of sentences. System accuracy can be enhanced based on various feature sets. Undergoing research shows that several feature extraction methods in both domains are promising [5]. Finally, future studies can also focus on conducting a comparative analysis of the obtained accuracy rates of ArSLRSs and determining what is the primary cause of this increase.

Note

Initial version of this paper was selected from the proceedings of International Marmara Sciences Congress (IMASCON 2020) which was held on June 19-20, 2020, and was subjected to peer-review process before its publication.

References

[1] M. Abdel-Fattah, "Arabic Sign Language: A Perspective", Journal of Deaf Studies and Deaf Education, vol. 10, no. 2, pp. 212-221, 2005. Available: 10.1093/deafed/eni007.

[2] Al-Ahdal, M. Ebrahim, and Md Tahir Nooritawati, Review in sign language recognition systems. In 2012 IEEE Symposium on Computers & Informatics (ISCI), pp. 52-57. IEEE, 2012.

[3] O. Al-Jarrah and A. Halawani, "Recognition of gestures in Arabic sign language using neuro-fuzzy systems", Artificial Intelligence, vol. 133, no. 1-2, pp. 117-138, 2001. Available: 10.1016/s0004-3702(01)00141-2.

[4] A., Alnahhas, B., Alkhatib, N., Al-Boukaee, N., Alhakim, O., Alzabibi and N., Ajalyakeen , "Deep Learning based Dynamic Hand Gesture Recognition with Leap Motion Controller", International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 5, pp. 7309-7315, 2020. Available: 10.30534/ijatcse/2020/61952020.

[5] M. AL-Rousan, K. Assaleh and A. Tala'a, "Video-based signer-independent Arabic sign language recognition using hidden Markov models", Applied Soft Computing, vol. 9, no. 3, pp. 990-999, 2009. Available: 10.1016/j.asoc.2009.01.002.

[6] A.S., Al-Shamayleh, R., Ahmad, N., Jomhari, and M.A., Abushariah, "Automatic Arabic Sign Language Recognition: A Review, Taxonomy, Open Challenges, Research Roadmap and Future Directions", Malaysian Journal of Computer Science, vol. 33, no.4, pp.306-343, 2020.

[7] R., Alzohairi, R., Alghonaim, W., Alshehri, S., Aloqeely, M., Alzaidan and O., Bchir, "Image based Arabic sign language recognition system", International Journal of Advanced Computer Science and Applications (IJACSA), vol. 9, no. 3, 2018.

[8] K., Assaleh, T., Shanableh, M., Fanaswala, F., Amin and H., Bajaj, "Continuous Arabic sign language recognition in user dependent mode", 2010.

[9] M.J., Cheok, Z., Omar and M.H., Jaward, "A review of hand gesture and sign language recognition techniques", International Journal of Machine Learning and Cybernetics, vol. 10 no. 1, pp.131-153, 2019.

[10] A.S., Elons, M., Abull-Ela and M.F., Tolba, "A proposed PCNN features quality optimization technique for pose-invariant 3D Arabic sign language recognition", Applied Soft Computing, vol. 13, no.4, pp.1646-1660, 2013.

[11] A.S., Elons, M., Abull-ela and M.F., Tolba, "Neutralizing lighting non-homogeneity and background size in PCNN image signature for Arabic Sign Language recognition", Neural Computing and Applications, vol. 22, no. 1, pp.47-53, 2013.

[12] S.M. Halawani and A.B., Zaitun, "An avatar based translation system from arabic speech to arabic sign language for deaf people", International Journal of Information Science and Education, vol. 2, no. 1,, pp.13-20, 2012.

[13] B., Hisham and A., Hamouda, "Arabic sign language recognition using Ada-Boosting based on a leap motion controller", International Journal of Information Technology, pp.1-14, 2020.

[14] N.B., Ibrahim, M.M., Selim and H.H., Zayed, "An automatic arabic sign language recognition system (ArSLRS)", Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 4, pp.470-477, 2018.

[15] M.M., Kamruzzaman, "Arabic Sign Language Recognition and Generating Arabic Speech Using Convolutional Neural Network", Wireless Communications and Mobile Computing, 2020.

[16] G., Latif, N., Mohammad, R., AlKhalaf, R., AlKhalaf, J., Alghazo and M., Khan, "An Automatic Arabic Sign Language Recognition System based on Deep CNN: An Assistive System for the Deaf and Hard of Hearing", International Journal of Computing and Digital Systems, vol. 9, no. 4,, pp. 715-724, 2020.

[17] H., Luqman, and S.A., Mahmoud, "Transform-based Arabic sign language recognition", Procedia Computer Science, vol. 117, pp.2-9, 2017.

[18] Miniwatts Marketing Group, "Internet World Users by Language", 2020. https://www.internetworldstats.com/stats7.htm, updated 11 Nov 2019. [Accessed: 11-Feb-2021]

[19] M. M., Mohamed, "Automatic system for Arabic sign language recognition and translation to spoken one", International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 4, 7140–7148, 2020.

[20] M., Mohandes, M., Deriche, U., Johar and S., Ilyas, "A signer-independent Arabic Sign Language recognition system using face detection, geometric features, and a Hidden Markov Model", Computers & Electrical Engineering, vol. 38, no. 2, pp. 422-433, 2012.

[21] J., Murray, "World Federation of the deaf", 2018. Rome, Italy. Retrieved from http://wfdeaf.org/our-work/. [Accessed: 11-Feb-2021].

[22] M., Mustafa, "A study on Arabic sign language recognition for differently abled using advanced machine learning classifiers", Journal of Ambient Intelligence and Humanized Computing, pp.1-15, 2020.

[23] M., Mustafa, H. , AbdAlla and H., Suleman, "Current approaches in Arabic IR: A survey", in International Conference on Asian Digital Libraries, pp. 406-407, 2008, December. Springer, Berlin, Heidelberg.

[24] M., Mustafa, A.S., Eldeen, S., Bani-Ahmad and A.O., Elfaki, "A comparative survey on arabic stemming: approaches and challenges", Intelligent Information Management, vol. 9, no. 2, p. 39, 2017.

[25] S., Reshna, and M., Jayaraju, "Indian Sign Language Recognition System–A Review", in International Conference on Signal and Speech Processing, ICSSP, vol. 14, 2014.

[26] Y., Saleh, and G., Issa, “Arabic Sign Language Recognition through Deep Neural Networks Fine-Tuning”, 2020.

[27] T., Shanableh, and K., Assaleh, “User-independent recognition of Arabic sign language for facilitating communication with the deaf community”, Digital Signal Processing, vol. 21, no.4, pp.535-542, 2011.

[28] V., Sharma, V., Kumar, S.C., Masaguppi, M.N., Suma and D.R., Ambika, “Virtual talk for deaf, mute, blind and normal humans”, in 2013 Texas Instruments India Educators' Conference, pp. 316-320, 2013, April. IEEE.

[29] P., Shukla, A., Garg, K., Sharma, and A., Mittal, “A DTW and Fourier Descriptor based approach for Indian Sign Language recognition”, in 2015 Third International Conference on Image Information Processing (ICIIP), pp. 113-118, 2015, December. IEEE.

[30] M.F., Tolba, M.S., Abdellwahab, M., Aboul-Ela, and A., Samir, “Image signature improving by PCNN for Arabic sign language recognition”, Canadian Journal on Artificial Intelligence, Machine Learning & Pattern Recognition, vol. 1, no. 1, pp.1-6, 2010.

[31] M.F., Tolba, A., Samir and M., Aboul-Ela, “Arabic sign language continuous sentences recognition using PCNN and graph matching”, Neural Computing and Applications, vol. 23, no. 3-4, pp. 999-1010, 2013..

[32] A., Wadhawan and P., Kumar, “Sign Language Recognition Systems: A Decade Systematic Literature Review”, Archives of Computational Methods in Engineering, pp.1-29, 2019.

[33] S., Wei, X., Chen, X., Yang, S. Cao, and X., Zhang, “A component-based vocabulary-extensible sign language gesture recognition framework”, Sensors, vol. 16, no. 4, p. 556, 2016.