Advances in machine translation for sign language: approaches, limitations, and challenges

Article *in* Neural Computing and Applications · May 2021

DOI: 10.1007/s00521-021-06079-3(

6 authors, including:

Mohd Shafry Mohd Rahim
Universiti Teknologi Malaysia

Nabeel Sabir Khan
University of Management and Technology (Pakistan)

Adnan Abid
University of Management and Technology (Pakistan)

Neural Computing and Applications
https://doi.org/10.1007/s00521-021-06079-3

ORIGINAL ARTICLE

Advances in machine translation for sign language: approaches, limitations, and challenges

Uzma Farooq[1,2,4] · Mohd Shafry Mohd Rahim[1] · Nabeel Sabir[2] · Amir Hussain[3] · Adnan Abid[2]

Received: 17 June 2020 / Accepted: 20 April 2021

Abstract

Sign languages are used by the deaf community around the globe to communicate with one another. These are gesture-based languages where a deaf person performs gestures using hands and facial expressions. Every gesture represents a word or a phrase in the natural language. There are more than 200 different sign languages in the world. In order to facilitate the learning of sign languages by the deaf community, researchers have compiled sign language repositories comprising of gestures. Similarly, algorithms have been proposed to translate the natural language into sign language, which is subsequently converted into gestures using avatar technology. On the other hand, several different approaches for gesture recognition have also been proposed in the literature, many of which use specialized hardware. Similarly, cell phone applications have been developed for learning and translation of sign languages. This article presents a systematic literature review of these multidisciplinary aspects of sign language translation. It provides a detailed analysis of carefully selected 147 high-quality research articles and books related to the subject matter. Specifically, it categorizes different approaches used for each component, discusses their theoretical foundations, and provides a comparative analysis of the proposed approaches. Lastly, open research challenges and future directions for each facet of the sign language translation problem have been discussed. To the best of our knowledge, this is the first comprehensive survey on sign language translation that discusses state-of-the-art research from multi-disciplinary perspectives.

Keywords Sign language · Natural to sign language translation · Gesture recognition · Avatar technology · Sign language repositories

1 Introduction

According to the World Health Organization, 5% of the world's population is suffering from hearing disability. Though this number seems small, it actually means that more than 460 million people around the world are affected by hearing loss. Out of these, 34 million are children. Furthermore, it is expected that by 2050, more than 900 million people will suffer from hearing loss [8], while 1.1 billion young people are at risk of hearing loss due to exposure to noise and other related problems. Unaddressed hearing loss results in a global cost of 750 billion US dollars [8]. The hearing loss is categorized into *mild*, *moderate*, *severe*, and *profound* categories based on the intensity of the deafness. The people with severe or profound hearing loss are unable to listen to the others and thus face problems in communication. This lack of communication can have a strong impact on the personality of the deaf person which may include loneliness, isolation, and frustration.

The deaf community uses a gestures-based language for communication, generally referred to as, sign language. Deaf people use sign language gestures to communicate with each other using specific signs. However, the hearing community does not understand these gestures, which

✉ Adnan Abid
adnan.abid@umt.edu.pk

1 Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Malaysia

2 Department of Computer Science, University of Management and Technology, Lahore 54000, Pakistan

3 Edinburgh Napier University, Edinburgh, UK

4 Department of Software Engineering, University of Management and Technology, Lahore 54000, Pakistan

poses a barrier for the communication between a deaf and a hearing person. There are nearly 200 sign languages in the world, and like the spoken languages, sign languages are also different from one another.

The developed countries have been facilitating their deaf community by providing them with tools and technologies to compensate for their cognitive and sensory loss. The people from hearing disability can be helped in many ways. Medically, an early diagnosis, use of hearing aids, cochlear implants are used as supporting means for the deaf community. On the other hand, assistive technologies, captioning the sign language, and other ways of educational and social supports are also used to address this problem. Apart from this, information technology-based solutions are also being developed to facilitate the integration of the deaf community in the mainstream. These solutions involve dictionaries and repositories to learn a sign language, translation systems that aim to convert voice or text into equivalent sign language gestures using avatar technology, and to close the loop, recognize gestures and produce natural language text or voice.

1.1 Motivation

Machine translation is one of the main research areas in natural language processing domain. Researchers have been working in this area for many decades. The surveys on machine translation approaches reveal that some initial work in this area was conducted in 1940s [23]. Ever since this area is evolving and many different approaches, data sets, and applications have been developed for the translation of one language into another language. There exist review and survey articles that cover approaches for natural language translation, in general [128], or for a specific set of languages [23, 65].

Sign language translation also falls under the umbrella of natural language translation problems. However, the specialized requirements and clear differences from natural language make it a significantly different branch of machine translation problems. On the one hand, it involves natural to sign language translation whereby text or voice in a natural language is converted into corresponding sign language gestures, while on the other hand there are gesture recognition approaches that aim to translate gestures performed by a deaf person into equivalent natural language text.

There exist surveys on both directions of translation. In [68], researchers discuss the natural to sign language translation approaches and projects. This survey discusses relevant details of prominent European and American projects conducted on natural to sign language translation systems. But it ignores the discussion on sign writing notations and data sets used for this purpose. Another relevant review study [80] presents a critique on natural to sign language translation project for ASL conducted during 1990s and early years of the decade 2000.

Similarly, some surveys for gesture recognition approaches have also been published. Some of these studies cover a subset of approaches that use hardware equipment to recognize human gestures [50, 112], or sign language gestures [59], whereas another survey discusses and compares the approaches that make the usage of Microsoft (MS) Kinect, different types of sensors, and gloves to recognize sign language gestures. Apart from this, several approaches exist that use computer vision-based approaches to recognize human gestures, which are discussed in detail in Sect. 6.

A recent study [31] highlights the important aspects involved in sign language generation, recognition, and translation. The authors highlight the importance and relevance of interdisciplinary research for sign language translation problem. They mainly emphasizes on cross-disciplinary research and presents promising future directions. However, it does not present a review and comparison of existing approaches for sign language translation.

This study presents a comprehensive survey of various different aspects involved in sign language translation problem. These aspects include: (i) sign language repositories and dictionaries [62, 111, 117]; (ii) natural to sign language translation [33, 80, 92, 134]; (iii) gesture recognition approaches [45, 64, 97, 139, 146, 190]; (iv) avatar technology for gesture generation [30, 37, 51, 55, 155]; and (v) applications for sign language learning and translation [37, 62, 117, 125]. Apart from this, future research directions for all areas have also been presented. This survey discusses the fundamentals of sign languages and then discusses and compares the research work done in all areas of sign language translation. To the best of our knowledge, this is the first comprehensive systematic literature review on sign language translation that not only discusses state of the art w.r.t. aforementioned aspects of sign language translation, but also presents a taxonomy of sign language translation and future research directions for each facet of this multi-disciplinary research area.

1.2 Structure of the article

The rest of the article has been structured in the following manner. Section 2 presents the fundamental details related to the sign language. Section 3 presents the research methodology adopted for conducting this survey which includes the research questions, inclusion and exclusion criteria, searching strategy, and quality assessment of the included studies. Section 4 discusses the details about sign language repositories and sign writing notations. Section 5 presents a detailed discussion on natural to sign language translation, while the research work related to the gesture recognition techniques is discussed in Sect. 6. Developments related to avatar technology are presented in Sect. 7. Furthermore, the applications related to sign language learning and translation are presented in Sect. 8. Subsequently, a taxonomy of sign language translation and open research directions are presented and discussed in Sect. 9. Lastly, the article is concluded in Sect. 10.

2 Fundamentals of sign languages

Sign language is a special type of language that is used for deaf individuals as their mode of communication. Unlike other natural languages, it makes use of meaningful body movements to convey the messages, and these body movements are called gestures or signs. Hands and finger movements, head nodding, shoulder's movements, and facial expressions are used to convey meaning. It is used by the deaf people for communication between deaf–deaf or deaf–normal individuals. When a hearing impaired individual wants to say something, he/she performs some gestures to communicate. Every particular sign means a distinct letter, word, or expression. Combination of signs makes a sentence just like words in spoken languages make sentences. Therefore, a sign language is a complete natural language with its own syntax and grammar.

Spoken languages vary from one region to another region, and about 6909 [182] spoken languages exist in the world. Similarly, there is no universal sign language, and nearly 200 different sign languages exist in the world. Every country has a different sign language, and some of them with their acronyms are presented in Table 1.

Table 1 Sign languages and their acronyms

Sign language	Acronym
American Sign Language	ASL
Arabic Sign Language	ArSL
Argentinian Sign Language	ArgSL
Australian Sign Language	AusLan
British Sign Language	BSL
Brazilian Sign Language	LSB
Chinese Sign Language	CSL
Greek Sign Language	GSL
German Sign Language	DGS
Indian Sign Language	ISL
Irish Sign Language	IrSL
Japanese Sign Language	JSL
Malaysian Sign Language	MSL
Mexican Sign Language	MxSL
New Zealand Sign Language	NzSL
Pakistan Sign Language	PSL
Portuguese Sign Language	PorSL
Russian Sign Language	RSL
Spanish Sign Language	LSE
Turkish Sign Language	TSL

2.1 Components of sign language

A sign language gesture involves two types of features, namely manual features and non-manual features. The manual features depend upon the shape, movement, location, and orientation of the hand. There are gestures which are performed by one hand, while the others are performed by involving both hands. Figure 1 highlights the manual features with the help of some gestures.

Non-manual features (NMF): Non-manual features include different facial expressions, head tilting/nodding, shoulder raising, mouthing, and related actions which adds meaning to our performed gesture/sign. Mostly, non-manual markers are used along with manual markers. Figure 2 highlights some non-manual features with the help of suitable gestures, while Fig. 3 presents structural differences between manual and non-manual gestures.

The gestures that involve hand movements are referred to as dynamic signs, while the gestures that do not involve any hand movement are termed as static sign gestures. Similarly, the gestures that involve both hands are called doubled-handed gestures, and the ones which are performed by a single hand are called single-handed gestures. Figure 4 presents some suitable gestures that help understanding the concepts of single-handed and double-handed static and dynamic gestures, where the dynamic gestures have been shown by presenting multiple frames. Categorization of gestures based on number of hands and movement is shown in Fig. 5.

Among the sign languages, American Sign Language (ASL) and British Sign Language (BSL) are based on English language, whereas Indian Sign Language (ISL) and Chinese Sign Language (CSL) are also among the well-known sign languages. Every country has its own sign language which means that for a particular given natural language word, there is a different gesture. Figure 6 shows gestures of different sign languages for the same natural language word. Furthermore, there exist different dialects of sign languages within the countries, especially in the big countries. The reason is that, in general, a sign language is taught by one person to another, and thus, there is a strong impact of localization on its creation, dissemination, and learning. Hence, there is a natural tendency that even within same country, there could be different signs for a given word. Almost every

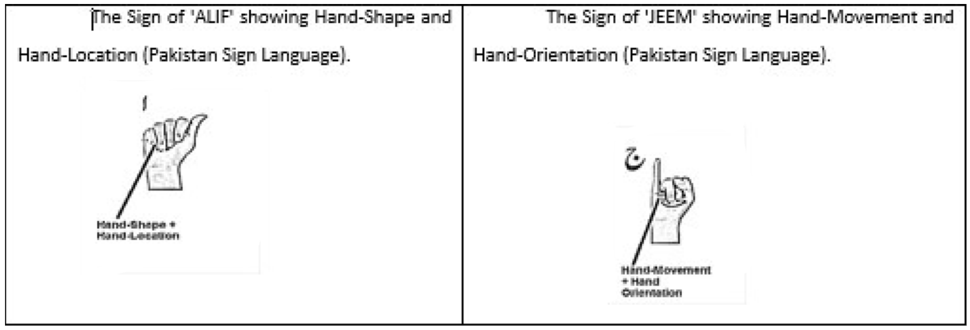


Fig. 1 Examples showing manual features

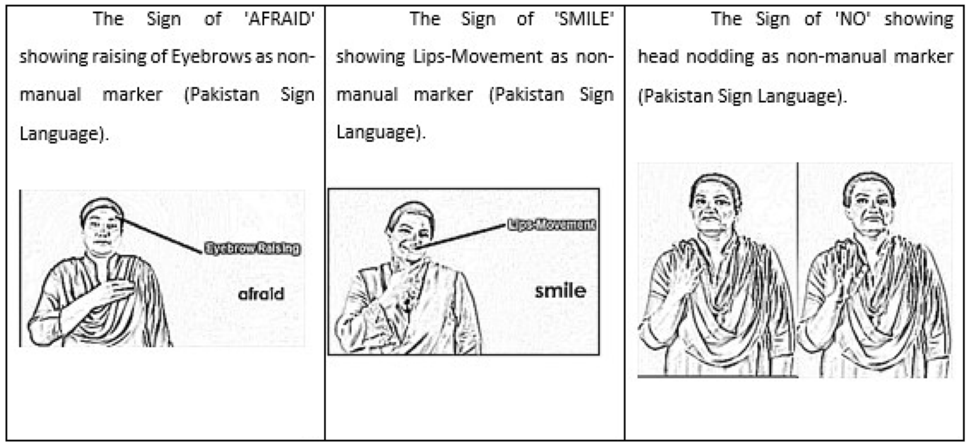


Fig. 2 Examples showing non-manual features

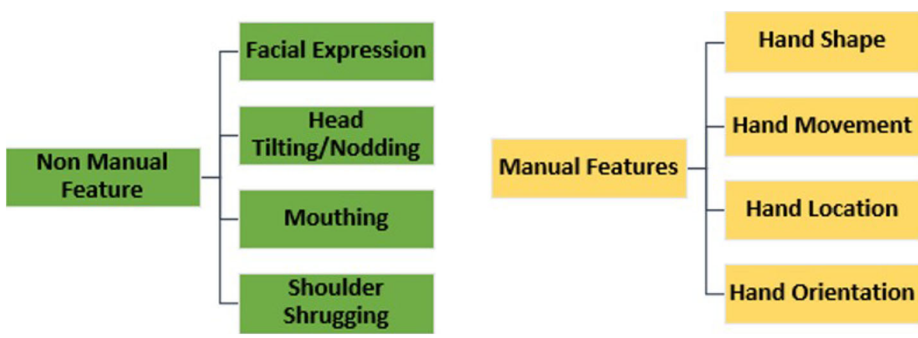


Fig. 3 Components of non-manual and manual features

Fig. 4 Single- and double-handed gestures

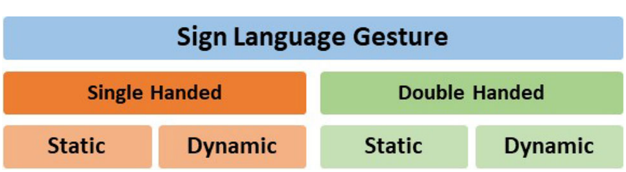


Fig. 5 Types of signs

Fig. 6 Different gestures for same word in different sign languages

sign language has one or more associated dictionary of gestures which are exposed on the Internet to learn and practice the language.

2.2 Linguistic information of sign languages

The grammars of these gesture-based sign languages differ from grammars of spoken or written languages [57, 92]. The reason is that gesture-based languages involve shapes and concepts, whereas spoken and written languages involve words and grammar rules, and thus, both types of languages have significantly different grammatical structures [48, 116].

Sign languages do not have well-defined word order, grammatical rules, and sentence structure. It does not have pre-defined standards of writing. Every sign language differs from the others in many terms like syntax, semantics, grammar, morphology, and phonology, whereby the syntax pertains to the rules that govern the arrangement of gestures to formulate a sentence. Semantics are responsible for validating the meaningfulness of the sentences. The structure of the language is defined by grammar. Similarly, the possible variants of formulating a sentence are catered for by the morphology.

Table 2 shows the grammatical different between English and Pakistan Sign Language. Simple sentences from

Table 2 Grammatical difference between English and PSL [92]

Category of sentence	English sentence and structure	PSL sentence and structure
Present indefinite	S + V + O	S + O + V
	Ali goes to school	Ali school go
Present continuous	S + HV + V(ing) + O	S + O + V + now
	Ali is going to school	Ali school go now
Present perfect	S + (has/have) + 3rd V + O	S + O + V(Ist form) + full
	Ali has gone to school	Ali school go full
Future indefinite	S + Neg + V	S + V + Neg
	Ali will go to school	Ali school go after
Past indefinite	S + adj + O	S + O + adj
	This is a correct algorithm	This algorithm correct
WH-interrogative	Wh + S + V	S + V + Wh
	Why Ali is going?	Ali going why?
Aux interrogative	Aux + S + V + O	S + O + V + Aux
	Do you want money?	You money want yes–no?
Compound	Sen + Conj + Sen	Sen + Conj + Sen
	Movie is good and seats are comfortable	Movie good and seats comfortable
WH-interrogative	Wh + S + V	S + V + Wh
	Why Ali is going?	Ali going why?

many different tenses have been presented. It can be clearly observed that the translated sentences in PSL are significantly changed in the following ways:

- The word order has been changed.
- Some words have been omitted
- Some new words have been added
- Words are in their lemmatized forms
- Prepositions and articles are not a part of translated sentence.

These differences highlight that natural to sign language translation is a complex problem. But at the same time, it can be seen that the deaf community cannot benefit from the translation of voice/text of natural language to textual form of equivalent sign language sentence. Rather the deaf community understands the gestural language.

Therefore, it is needed to transform these words into gestures of a sign language. As a matter of fact, the translated sentence helps in generating sequence of words whose gestures should be rendered one after the other to help the deaf person understand the meaning. It is pertinent to mention that the proper nouns like names of person, places, etc., are finger spelled by the deaf persons, which means that they are represented with the spellings of the letters that constitute a word, while for most of the words and phrases, there exist appropriate sign language gestures different languages.

Interestingly, the non-manual features also make a significant difference in the understandability of a gesture. For instance, in ASL, raised eyebrows represent an open-ended questions, while for a yes/no questions furrowed eyebrows are used. Likewise, mouth movement also plays its role; for instance, performing the gesture for the word "cup" with different mouth positions represents different sizes of a cup. Apart from this, the positioning of a subject/object in the sign space is also important, and the same subject/object can be referred to by finger pointing to the same point, or with the help of an eye gaze.

3 Methodology

This review has been conducted in a systematic manner as per the guidelines provided by Altman [106] and Kitchenham and Charters [96], whereby different research questions are defined to meet the objectives of the survey. The relevant scientific literature is found using a well-defined protocol so as to ensure the inclusion of important relevant research. The protocol includes: (a) defining appropriate search queries and applying them to appropriate research portals; (b) well-defined inclusion and exclusion criteria for choosing appropriate studies; and (c) method for study selection, extraction of useful information, and performing analysis of selected studies.

3.1 Defining research questions

The first step of the systematic review process was to define the research questions. Based on the aim of this study, the research questions presented in Table 3 were defined.

Table 3 Research questions and motivations

Questions	Motivations
1. How sign language repositories are being developed and managed?	Identify different methods for developing, maintaining, and extending sign language repositories
	Figure out the ways in which the gestures of a sign language can be digitized
2. Which methods have been developed to convert natural language into sign language?	Identify how the grammatical structure of a sign language is different from written natural languages
	To discuss and compare different machine translation models used for natural language to sign language translation
3. What are different methods to recognize the gestures of a sign language?	To identify different machine translation models developed for gesture recognition
4. How the deaf community is being supported by avatar technology to play automatic gestures?	To discuss the developments in avatar technology that is helpful for generating automatic gestures for different sign languages
5. What types of applications have been developed to facilitate the deaf community of the world?	Discuss various types of applications developed for different sign languages in the world
6. What are promising future research directions for the development of sign language technologies?	Discuss the major challenges for the development of tools and technologies in all the aforementioned components of sign language translation systems

3.2 Criteria for including and excluding studies

The following criteria were considered for inclusion of studies for this research:

- The context of the study proposes a solution to a problem related to the translation of sign language.
- The study addresses a particular dimension of the research questions. This means that the study should either focus on at least one of the following components of sign language translation: sign language repositories, translation from natural to sign language or vice versa, avatar technologies for sign languages, or applications related to sign language.
- The article should preferably be published in well-reputed conference or journal.
- The article is written in English language.

The exclusion criteria for the studies are the following:

- Any language translation model not addressing the sign language research has been excluded.
- Articles discussing the aspects other than sign language translation were excluded.
- Articles not written in English were excluded.

3.3 Search strategy

The articles were searched from Web of Science repository using the search string presented below.

(sign language) OR (sign language repository) OR (sign language dictionary) OR (sign writing) OR (sign writing notations) OR (sign language translation) AND (gesture recognition) AND (sign avatar) AND (sign language applications)

Figure 7 presents the step-by-step process for the selection of articles in this study. The initial search presented more than 1800 research articles. These articles were then scrutinized based on title and quick scan of their abstracts. During this process, only those articles were retained which address a sub-area of sign language translation including sign language repositories or dictionaries, sign writing notations, natural to sign language translation, gesture recognition approaches, avatar technology, and applications for sign languages. During the screening phase, more than 1600 articles were eliminated. Subsequently, to check the final eligibility, full text of the rest of the articles was explored, and 98 articles were eliminated at this stage for they were not directly related to the problem of sign language translation. This resulted into a set of 147 articles covering different aspects of sign language translation problem.

Figure 8a presents the distribution of selected articles w.r.t. the publication channels. It can be observed that more than 50% of the selected articles are published in journals, while almost 30% articles have been published in conference proceedings, and the rest of the references are from books, thesis, and web sites. One the other hand, Fig. 8b presents the proportion of articles belonging to a certain aspect of sign language translation. It can be observed that almost 40% of the studies discuss gesture recognition approaches, while 30% of the studies discuss natural to sign language translation, whereas 17% studies discuss about sign language dictionaries and repositories. Lastly, more than 10% references belong to applications based on

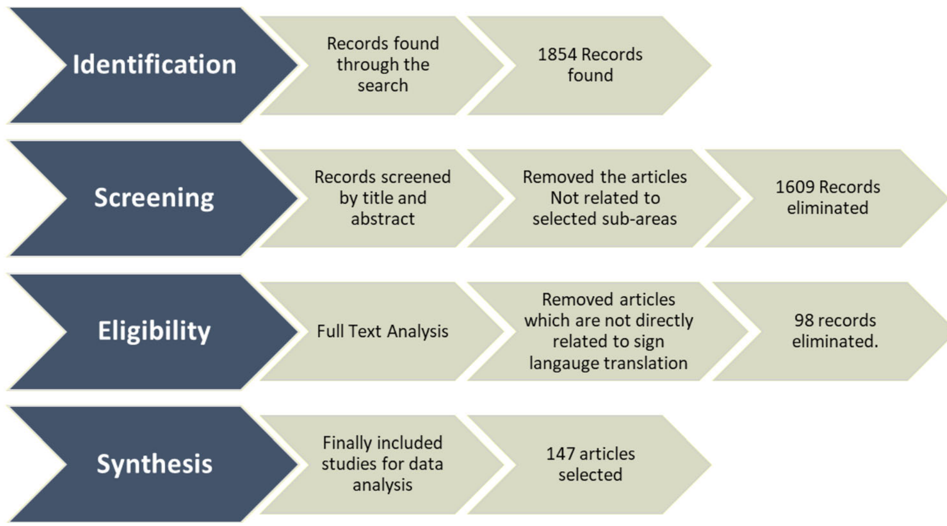


Fig. 7 Flowchart of systematic search process and selection of research studies

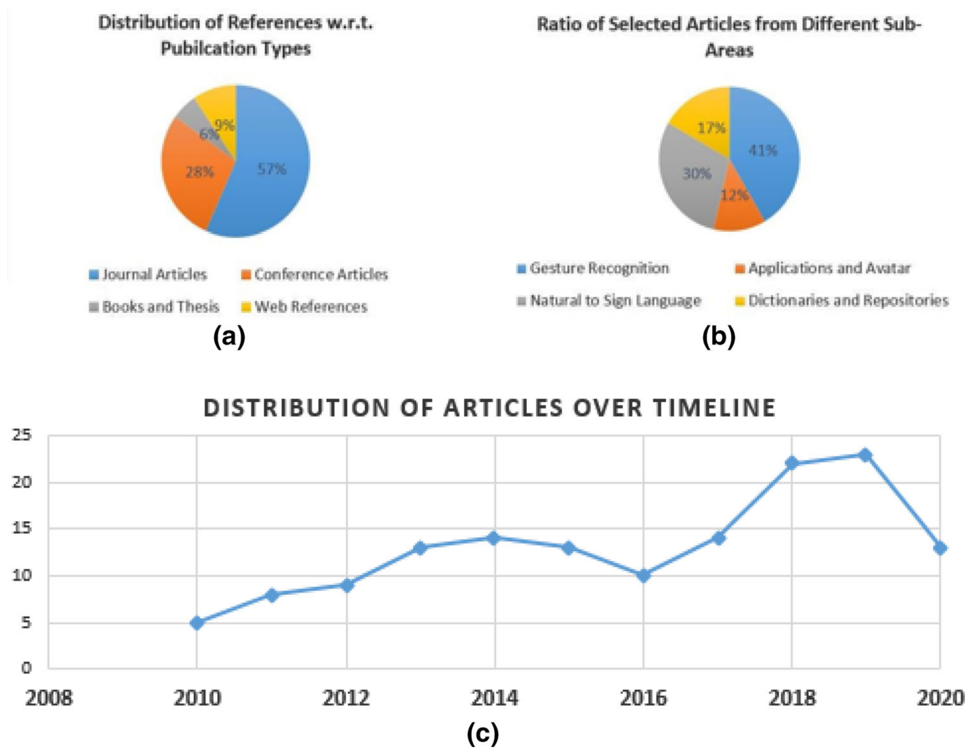


Fig. 8 **a** Distribution of references w.r.t. the publication channels. **b** Distribution of articles from different sub-domains of sign language translation. **c** Distribution of selected articles over timeline

sign language translation, and about the developments in the avatar technology for sign languages.

Figure 8c shows the distribution of selected articles over the timeline. Most of the selected articles fall in the bracket of 2015–2020. The data reflect that this survey not only contains high-quality articles published, but at the same time presents the synthesis of recently published articles in the area of sign language translation. More than 50% of the considered articles have been published after 2016, which ensures that the discussion in this survey is based on recent advancements in this area.

3.4 Quality assessment

Quality of publications has been a strong consideration while selecting the articles for formulating this review study. To this end, the journal articles have been selected based on their quartiles assigned by Scimago Journal Rank available at https://www.scimagojr.com/. This ranking system divides the journals into four quartiles Q1 to Q4, where Q1 includes the best journals, while the others have lower ranks accordingly. It can be observed from Fig. 9a that more than 50% of the selected article belong to Q1 category, while almost 75% of the selected journal articles

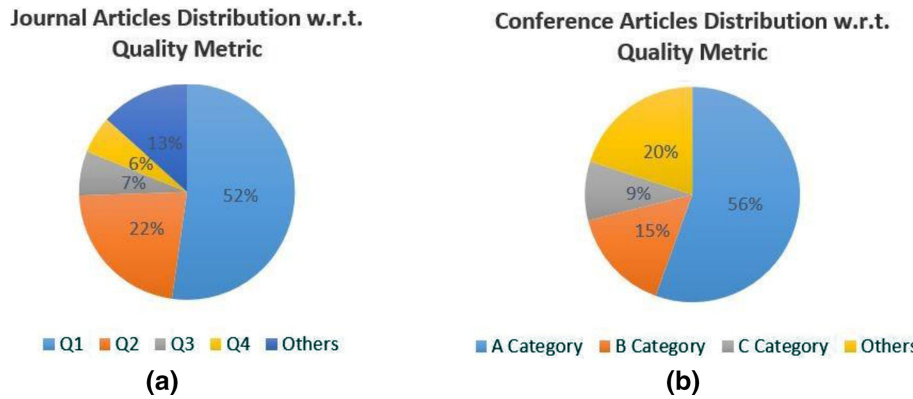


Fig. 9 Quality assessment of selected journal articles (**a**) and conference articles (**b**)

belong to Q1 and Q2 categories. On the other hand, the conference articles have also been selected very carefully while keeping the quality factor into consideration. The conferences are ranked by Conference Ranks that show a meta-ranking of conferences by involving CORE ranking maintained by Excellence in Research in Australia (ERA), and Microsoft Academic Research (MSAR). The conferences are categorized into A, B, and C categories, A being the best category, which contains most of the flagship conferences. It can be observed from Fig. 9b that more than 50% conference belong to A category conferences, while more than 70% of the conference articles belong to A and B categories. Thus, it has been ensured that the selected articles should possess excellent quality.

4 RQ1: sign language repositories and sign writing notations

4.1 Sign language repositories

Experts and communities from different countries have developed sign language repositories that largely comprise of sign language dictionaries. The repositories are also referred to as sign language dictionaries, corpus, and at times corpora. Most of these dictionaries comprise of most frequently daily used words in different useful multimedia formats. The most common representation in the sign language dictionaries is presentation of words in the form of pictures, videos, or in the form of avatars. In terms of understanding the gestures, the video representations are much easier as compared to the pictorial representations, while the avatar-based representations are also understandable, yet they are considered less expressive as compared to the videos. However, in terms of data storage the pictorial representation takes least amount of storage, followed by avatar videos, whereas the video representation takes the highest storage space. Some dictionaries have also stored the gestures in sign writing notations, which helps storing these gestures in a textual representation that can be later converted into a gesture using avatar technology. Obviously, being a text format it takes the least amount of storage, but requires tools to convert back into the equivalent sign language gesture. More details related to sign writing notations and avatar technology have been discussed in coming sections.

These dictionaries also vary in terms of granularity of language components. For instance, almost all of them include the gestures for all the letters and basic numbers for the considered sign language; the next granularity is that of storing the gestures for words, i.e., they store word of a spoken language along with its representation in the considered sign language. Many dictionaries also store some frequently used basic phrases, e.g., greetings, etc., which is another level of granularity. Furthermore, many languages have exposed the sign language-based children stories and poems which comprise of several sentences, which is yet another higher level of granularity. Most of these dictionaries present the information under different categories and are augmented with searching feature that helps finding the gestures of words easily. This study presents some details of dictionaries and repositories made and exposed by different sign language experts and organizations across the globe.

A British Sign Language corpus consisting of 2528 video clips was proposed in (2011) by [43]. These videos were recorded by the deaf people using BSL. The corpus is also exposed to public so that they can use and learn BSL. The European Cultural Heritage Online organization (ECHO) published corpora consisting of children stories and poetry in Swedish, British Netherlands SL. This corpus contains video signs performed by single signer. Similarly, [167] proposed a corpus for deaf children in Africa.

There are various multilingual dictionaries having different kinds of representations which are available for different SLs. Spanish Sign Language-Spanish (DILSE) dictionary was proposed by [63]. It is a multilingual dictionary which is available online for the deaf community of

Spain. The dictionary provides two levels of search in which user can either search a Spanish word against a sign and similarly can do a sign search by giving a Spanish word as input. The Italian sign language dictionary was proposed by [49]. An electronic multimedia dictionary consisting of signs of 3 different sign languages including American Sign Language (ASL), Japanese Sign Language (JSL), and Korean Sign Language (KSL) was created by [164]. A Danish Sign Language (DTS) dictionary was developed by [100]. The dictionary stores words along with their synonyms and corresponding human recorded videos of each sign.

Some groups worked on domain specific SL corpora, and details of some of the domain specific SL dictionaries are described here. One domain specific dictionary includes a weather reports corpus comprising of 2468 sentences in German and DGS and has been reported by [73]. An ISL Dictionary for disaster domain was proposed by [117]. The videos are traced to convert it into avatar animation. This dictionary provides the information of disaster to deaf. It consists of 600 sentences and 2000 words. Some dictionaries are enriched with sign writing notations, which help them in showing the gestures using avatar technology. For instance, a DGS Corpus was developed by [62] using Gloss and Hamburg Notation System (HamNoSys). Apart from this, very few repositories target the machine translation for sign languages. As an example, in [38] the researchers present statistical machine translation experiments on a corpus of about 2000 sentences for the language pair Chinese and French.

There are many sign language dictionaries available online. Repositories according to different regions along with total number of videos, format, URL, and sign writing notations are listed in Table 4.

It can be observed that most of the dictionaries have video representation of signs, but some contain an image as well. These dictionaries contain signs of words, and some of them, e.g., "lifeprint," also include signs of common phrases. There are also few YouTube channels like "Elma Production" that are uploading sign language lessons. Mobile application dictionaries are also available for some languages such as Pakistan Sign Language [4] and American Sign Language [15]. Similarly, majority of these repositories are not machine-readable and only contain videos and images. Following are the limitations of multimedia-based representation of sign language gestures which invite the researchers to devise methods to develop suitable sign writing notations (Table 5).

Storage size: The size of storing a gesture in image or video format requires substantially large amount of storage as compared to text-based storage. This challenges the scalability of large-scale systems.

Lack of representation in image-based gestures: The dynamic gestures are hard to represent through image-based storage, as they are presented through multiple frames showing different hand shapes and orientations, while the movements are shown with the help of different symbols like arrows, etc.

Standardization: It is hard to standardize the representation of images- and video-based gesture repositories in many ways. For example, engaging the same person for performing the gesture is not easy. Similarly, maintaining the same background settings, clothes, etc., are also challenging.

Issues in sentence generation: The gestures are connected with one another in a specific sequence to reflect a meaningful sentence. This can certainly be done by rendering the gestures of these words in given order. However, this poses a lot of challenges while using image- or video-based repositories. For example, gestures recorded by different persons in different settings or by same person with different clothes result in user experience of the translation system, while in case of image-based repository it becomes even more tedious where different frames are loaded while showing the arrow signs representing the movements. Figure 10 shows the generation of different sentences using image and video gesture repositories. It is clear that gestures for the words involving in a sentence have been performed by different persons, thus making it difficult to understand.

But there exist some online data sets [6, 7, 11, 14] that includes some basic signs videos and images for the validation of developed approaches. American Sign Language (ASL) includes graphemes and Gloss writing of signs. There is also an online guide for sign writing. Another important aspect of sign language dictionaries is the standardization of gestures for each language. Since, every region has its particular sign representation that varies from region to region, thus it is imperative to have a process to add new gestures in a language while involving standard bodies. To this end, almost every country has a responsible organization that is responsible for maintaining the standard sign language gestures, e.g., Americans National Institute of Deaf (NAD) is responsible for ASL [4]; Pakistan Directorate General of Special Education Initiative National Institute of special education(NISE) [9] is responsible for managing PSL, while Indian National Institute of the Deaf [13] does it for ISL.

4.2 Sign writing notations

Notational systems are significantly important in the automation of system [169]. All the natural languages can be written using their alphabets. Similarly, sign languages can also be written using different gesture writing systems.

Table 4 Sign language dictionaries list

Language	Total word	Format	URL	Sign writing notation
American Sign Language (ASL)	Over 2650 signs	Image/ Video	https://www.lifeprint.com/	
	Over 7835 signs	Video	https://www.handspeak.com/	Graphemes
	10 lessons	Video	http://www.signlanguage101.com/	
	Over 9273 signs	Video	https://www.signingsavvy.com	ASL Gloss
British Sign Language (BSL)	Over 21,000 signs	Video	https://www.signbsl.com/	
	Over 2500 signs	Video	https://bslsignbank.ucl.ac.uk/	
	Over 478 signs	Image	http://www.british-sign.co.uk	
German Sign Language (DGS)	Over 6000 signs	Video	www.dgs-korpus.de (in development phase)	
	Over 15,000 signs		https://www.spreadthesign.com	
Spanish Sign Language (LSE)	9 lessons containing multiple signs		http://www.cervantesvirtual.com/seccion/signos/	
	Over 15,000 signs		https://www.spreadthesign.com	
Indian Sign Language (ISL)	Over 2775 signs	Video	http://www.indiansignlanguage.org	
	Over 3000 signs	Video	https://www.youtube.com/channel/UC3AcGIlqVI4nJWCwHgHFXtg	
	Over 898 signs		https://www.talkinghands.co.in	
Pakistan Sign Language (PSL)	Over 7000 signs	Video	http://www.psl.org.pk	
Arabic Sign Language (ArSL)	Over 2861 signs	Video	http://sldictionary.appspot.com	
Chinese Sign Language (CSL)	Over 542 signs	Video	https://www.newsigns.jp/fsle/china	
	Over 44 signs	Video	http://www.deafstudies.jp	
Malaysian Sign Language (MSL)	13 lessons containing multiple signs	Video	https://www.youtube.com/user/ElmaProduction/feed	
		Video	https://books.google.com.pk/books/about/MySLang.html?id=C89NAQAACAAJ&redir esc=y	
Australian Sign Language (AUSLAN)	Over 7983 signs	Video	http://www.auslan.org.au/	
	Over 345 signs	Video	https://www.youtube.com/channel/UCjCBXZL4T0nCyq3AgFVsDg	
South African Sign Language (SASL)	Over 800 signs	Video	https://www.realsasl.com	
	Over 61 Videos having multiple signs	Video	https://www.youtube.com/channel/UCpuibTErIqB3oEvp0Jaap9w	
Brazilian Sign Language (LSB)	9500 entries	Images	http://www.signwriting.org/brazil/brazil21.html	

Though different sign writing notations systems exist, none of them has been considered to be a standard sign writing system. It is clear from the previous section that most of the sign language gesture repositories store and present the gestures in the form of videos or images. However, the sign writing notations help in preserving the gesture in a textual format.

Many different sign writing notation systems have been introduced by the researchers. Stokoe, Sign Writing, Gloss, and HamNoSys are widely used sign writing notations [81, 160].

4.2.1 Stokoe notation system

The Stoke Writing Notation System was the very first sign writing notation system introduced and developed by William Stokoe. As it was the base line defined for Sign Writing Notation Systems, many other notations are based

Table 5 Comparison of sign writing notations

Sign writing notation system	Sign language dependent	Non-manual features support	Objective	Arrangement	Computer compatibility
Stokoe	Yes	No	Dictionary or Academic	Linear	Custom Font or ASCII codes
Gloss	Yes	Yes	Academic	Linear	Custom Font or ASCII codes
SignWriting	No	Some	Public Use	Pictorial	ASCII or Unicode
HamNoSys	No	Yes	Academic	Linear	Custom Font Unicode

Fig. 10 Representation of gesture using images or videos

on Stokoe's Notation. Stokoe defined signs by introducing following three main parameters which he called "aspects" [81].

Hand configuration: Nineteen separate instructions or hand design shapes were defined by Stokoe, including open palm, closed fist, or partly closed fist with the index finger.

Place of articulation: It has twelve shapes which deals with the signs made at the upper arm, upper brow, or the cheek. Movement: It contained twenty-four values to determine whether the hands are up, down, side-by-side, forward and away, rotating, and so on. Few symbols are shown in Fig. 11.

4.2.2 Gloss notation system

In this notation system, signs are represented using words from English Language or some other language. The translation of signs in Gloss requires a dictionary-based approach [81]. An example expression is shown in Fig. 11.

4.2.3 Sign writing notation system

Sign Writing Notation System was introduced by Valerie Sutton who was a dancer. It was initially developed for communication between deaf people. Among all Sign Writing Notation Systems, sign writing is considered as practical one to be implemented. Few symbols are presented in Fig. 11.

4.2.4 HamNoSys sign writing notation system

It is essentially a collection of notations for Hamburg Sign Language developed in 1985 in Germany by the University of Hamburg [138, 155] and has its own pre-defined note and phonetic transcripts for sign and gesture definitions. It offers us the possibility to write signs in a computer that is easier to understand and process.

HamNoSys's origin was primarily a Stokoe notation system that gives us an alphabet to define different parameters for sign language, such as hand shape, manual movements, location of hands, and hand orientation. Few symbols are shown in Fig. 11.

Fig. 11 Different sign writing notations

The comprehensiveness of the sign writing notation is very important for its usage in translation system, where the representation of hand shapes, orientation, location, and movements is important. Similarly, the coverage of non-manual features is also important for complete representation of a gesture. A comparison of these notations systems showing their characteristics and strengths is presented in Table 6. Among these sign writing notations, Gloss and HamNoSys offer the coverage of non-manual features [81]. However, HamNoSys offers much more external support as there exist keyboards, tools, and services that help converting a gesture writing in HamNoSys into equivalent animates character [3, 26]. Furthermore, HamNoSys is not dependent on any sign language, rather can be used to write any arbitrary hand movements. The storage format and linear representation reflect that it can be written in a textual format which results into reduced storage cost. Hence, HamNoSys is the most frequently used sign writing notation by the researchers and practitioners. Figure 12 presents the details as to how HamNoSys covers different features of a gesture in its vector representation.

4.3 Advantages of sign writing notations

It can be anticipated that writing a gesture in appropriate sign writing notation may offer many advantages, flexibilities, and extensibilities while using them for gesture

Table 6 Comparison of image, video, and sign writing notations-based gesture repositories

Storage type	Pros	Cons
Video signs	Close to reality	Requires standard recording settings
	Easy to create with the help	Requires considerable storage space
	of sign language expert	Not suitable for sentence level translation systems
Images/pictures	Lesser memory requirements as compared to videos	Difficult to create
		Requires sign language experts
		Less realistic as compared to videos
		Not suitable for sentence level translation systems
Sign writing notations	Take least memory storage	Animated character is not as good as real person's gesture
	Can be rendered seamlessly to generate sentence level translation	Have to write gestures
		manually (so far no automatic sign writing systems exist)

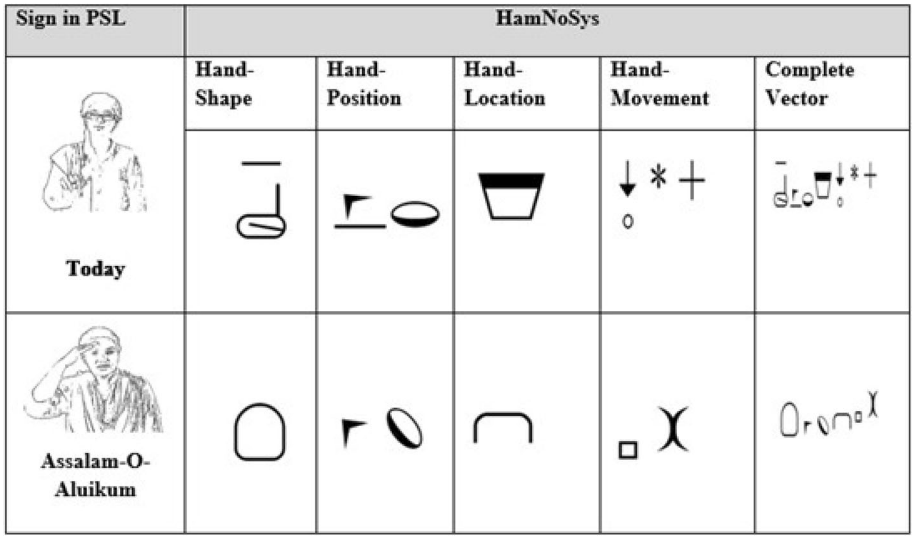


Fig. 12 Writing HamNoSys vectors for gestures

generation in natural to sign language translation systems. Some of the salient advantage are the following:

Offer reduced storage:

The sign writing notations require significantly less secondary storage to store the gestures' vocabulary. Similarly, at the time of rendering a gesture in machine-readable format, very little portion of main memory is required as compared to loading a multimedia image or video.

Can generate gesture of a sign dynamically:

Storing a gesture in machine-readable format provides the flexibility to render it and generate avatar using different suitable characters for different age groups. Similarly, the background environment and settings can be changed as per requirements. Thus, it can help building a dynamic and standardized translation system, where a typical character can perform animated gestures in a standardized settings, thus giving a better user experience.

Figure 13 shows the animated gestures for the sentence "Mother is driving a car." The gestures are played in a video, and it can be seen that a single avatar is performing all the gestures in given settings. This resolves the problem shown in Fig. 10.

Offer seamless extension:

The gesture repository can be extended in a seamless manner in the form of machine-readable format. As otherwise, it is hard to engage the same person again in the same settings to record videos or capture images. The gestures are gathered over the years, and new gestures are also generated for newer terms and words. Thus, even if we engage the same person, the settings and orientation of the person may not persist. Similarly, new gestures can be seamlessly added using sign writing notations and can be exposed with the help of different visualization technologies.

Benefit from evolving technology:

The gestures can be translated to newer and more comprehensive writing notations (though this requires extra effort). Similarly, better tools and technologies that convert the sign writing notations into gestures can be used at any stage.

It is clear that there are several advantages and disadvantages of each representation. A comparison of three different formats for storing the gestures is presented in Table 6.

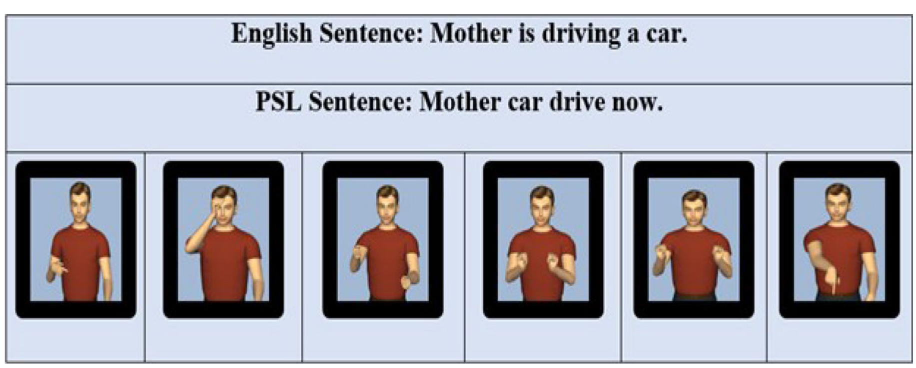


Fig. 13 Video generation using avatar

5 RQ2: natural to sign language translation systems

There are several different models and approaches for machine translation. Statistical machine translation, semantic machine translation, and neural machine translations and their wider categories are shown in Fig. 14.

Every natural language uses many alphabets and words, and from those words, different sentences are formed. It is pertinent to mention that the sentence structure of a sign language is very different from that of natural language. Table 2 highlights the differences between the sentence structure of natural and Pakistan sign language sentences.

All three MT approaches, namely semantic approaches, statistical approaches, and neural machine translation-based approaches, have been used for sign language translation problems. It has been observed that most of the approaches are based on transfer-based or interlingua-based semantic approaches [23, 65]. Very few of them involve statistical machine translation, while recently, there are a couple of efforts to use deep learning-based approaches for natural to sign language translation.

5.1 Semantic machine translation

Semantic translation models are used for machine translation as well, and they rely on text processing using syntactic and semantic analysis while using some rules or ontologies. These systems emerged in 1990s. These systems exhibit reasonable accuracy, but are limited to the rule base or ontology.

In semantic translation model, three variants, namely direct translation, interlingua and transfer based, exist for

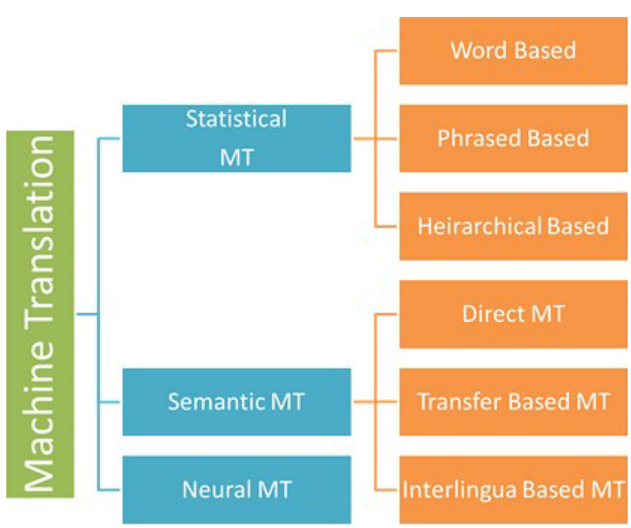


Fig. 14 Classification of machine translation approaches

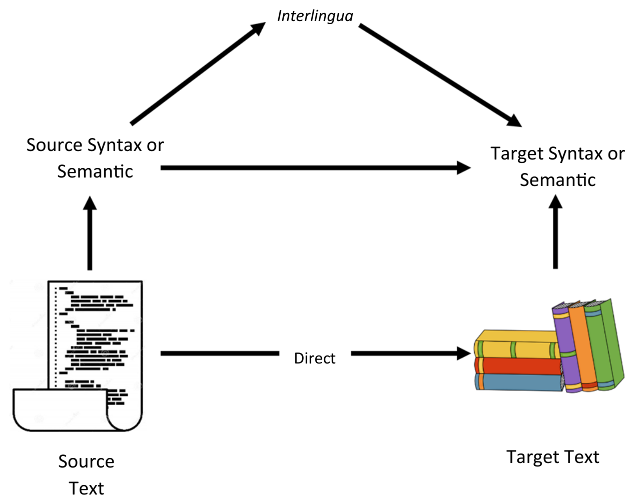


Fig. 15 Semantic machine translation model

the conversion of translation of natural language, as shown in Fig. 15.

5.1.1 Direct machine translation

It is the most basic form of MT. In this approach, the words from source language are replaced with the corresponding word from the target language. Such systems process each word of the source language on individual basis. Conversion of the input sentence is done without any type of syntactic and semantic analysis on the source input sentences as shown in Fig. 15. It is important to note that the sequence of words remains the same after the translation, whereas the sequence of words in sign languages is not the same as that of the written language. Therefore, the direct translation approach is not suitable for translation of source and target languages having different sentence structures and, hence, is not suitable for sign language translation systems.

Such systems process each word of the source language on individual basis. Conversion of the input sentence is done without any type of syntactic and semantic analysis on the source input sentence. TESSA and SignSynth projects are the applications of direct translation systems.

TESSA: This is essentially a British Sign Language (BSL) interpretation system that supports the correspondence between a post office staff member and a person who is hearing impaired. This system uses the English term as an input phrase, searches the English to sign dictionary with each letter of the input text, puts identified signs together, and executes their animation for each word. This system follows direct translation approach [44].

SignSynth project: This system used ASCII-Stokoe model to represent signs. The result is produced in animation format by transforming ASCII-Stokoe into VRML (Virtual Reality Modeling Language) [69].

Another project of machine translation system to a text to American Sign Language (ASL) was proposed and developed which was restricted to the domain of weather information [69].

5.1.2 Transfer-based machine translation:

This level lies in between the direct translation and interlingua translation systems, and most translation systems fall into this category. The initial intent of transfer approach is to apply appropriate transformations so as to generate a syntactically correct text of the target language. Although the rules used for the transfer depend on both the source and target languages, some of the rules may need only slight modification when a MT system is developed for a new target language linguistically related to an existing one. Note that for applying transfer approach, we need some "linking rules" that map between the source language text and target language text.

Such systems analyze and process the input sentence either syntactic or semantic to some level. For such conversions, a proper set of conversion rules are designed that reads the source text and employs semantic or syntactic structure in the respective targeted sign language. TEAM, ASL workbench, and ViSiCAST translator are the example of such systems that work on the principle transfer-based translation systems.

TEAM: This assignment is a text to American Sign Language (ASL) translation system. It has the Synchronous Tree Adjoining Grammar (STAG) which was developed for the representation of input text into its respective ASL syntactic structure. It has a component of bilingual lexicon in order to identify and verify the valid and correct word-sign pair [192].

ASL workbench: A machine translation system converts input text to American Sign Language (ASL). It used Lexical Functional Grammar (LFG) for the representation of English f-structure in ASL. Proposed system used a reliable and efficient phonological model that was based on Movement-Hold principle of American Sign Language phonology [46].

ViSiCAST translator: A tool which converts English input sentence to British Sign Language (BSL). The system used HPSG for the representation of input text in BSL. The grammar part was designed and developed using a Prolog-based system ALE. BSL phonology was represented in HamNoSys sign writing notation system [54, 55].

5.1.3 Interlingua machine translation

These systems semantically process and analyze the input source phrase to automate an independent linguistic semantic structure called an Interlingua, and finally, a generation component gives the output phrase into its respective sign language. It is an expensive process because the semantics of the source language must be preserved.

This is considered to be the most degenerate form of transfer, i.e., source language text should be translated into an abstract and language-independent representation, generally known as an interlingua. Subsequently, the target language sentence is generated from this intermediate representation as shown in Fig. 15. It is a challenging process in terms of maintaining the semantics of the source language.

Following are interlingua-based semantic machine translation systems for sign language translation:

ZARDOZ: This system is an example of interlingua-based translation system. A system was proposed which translates English input text into its respective sign language. Indian Sign Language (ISL) and Japanese Sign Language (JSL) are the targeted sign languages. As an Interlingua for expression, a variety of hand-coded systems have been used. The system takes English text as input and according to the choice of the user translates it either in JSL or in ISL [172].

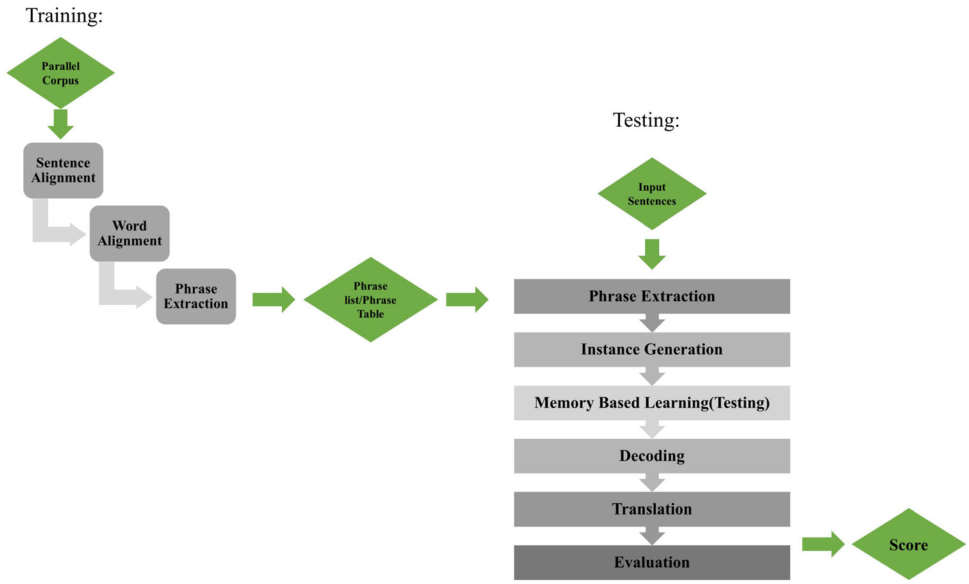
5.1.4 Recent research using semantic approaches

Apart from the aforementioned classical projects, some recent semantic translation systems for European and Asian Sign Languages have been developed recently, for instance, Greek Sign Language [99], Spanish Sign Language [135], Arabic Sign Language [113], Vietnamese Sign Language [123], Indian Sign Language [173], Bangla Sign Language [58], and Thai Sign Language [171]. These systems are based on transfer-based and interlingua-based translations, which rely on rules for machine translation.

5.2 Statistical machine translation

Statistical machine translation is also referred to as machine learning-based translation method [107]. In this approach, it is assumed that with some probability score, each word in target language is a translation of the source language word. The best translation is determined by a sequence of words which has highest probability.

Fig. 16 Workflow for statistical machine translation systems [157]



The main step in statistical machine translation is corpus preparation, training, decoding, and testing as shown in Fig. 16, whereby the pre-processing phase involves corpus preparation, cleaning, and alignment of sentences, where alignment plays important role in training the corpus. Decoding is a major phase in which target language sentences are being decoded [157].

Two major challenges in this approach are reordering of language sentences and complexity of decoding the sentences. Further this translation model has three granularity levels of application namely: word-based, phrase-based, and hierarchical-based statistical machine translation [23].

In word-based translation, word-to-word translation is required with some alignment which follows certain patterns. This is an oversimplified approach that negatively impacts the quality of translated output, whereas in phrase-based approach the sentences of source and target language are divided into separate phrases. Alignment is done using reordering the patterns. This approach works well for the small sentences, but does not perform well with long and complex sentences, while the third variant is hierarchical phrase-based model, which is better than previous two approaches as it has recursive structure instead of simple phrases that leads to improve the accuracy of SMT system.

There exists some research work on natural to sign language translation using statistical machine translation approaches. A couple of such models have been proposed for ASL to gloss by [28, 127], where [28] has used Moses toolkit system for development and evaluation and has achieved BLEU score of 0.51 without turning and 0.62 after turning.

Similarly, statistical machine translation systems for some European Sign Languages text to their respective sign languages have also been developed. For example, [149] deals with Spanish Sign Language, while [101] translated Czech to Czech Sign Language. A workflow that involves semantic/rule-based and statistical translation models is presented in Fig. 17.

On the other hand, the statistical machine translation methods have mostly been used for gesture recognition rather than natural to sign language translation, e.g., some recent work has been done on Arabic [82, 114], Indian [90], and Malaysian [90] sign languages' gesture recognition.

5.3 Neural machine translation

Machine and deep learning-based methods are widely being used in natural language processing applications [130, 141, 177]. The most recent form of machine translation involves neural machine translation models. These models use neural networks to translate source text into target text. With the increasing amount of data and availability of sophisticated hardware, deep learning methods are gaining more popularity for machine translation tasks as well [187]. Since this study aims to involve deep learning which falls under the area of neural machine translation, therefore relevant details for this area have been discussed in detail.

Deep learning: Deep learning is a specialized form of machine learning that manages to solve problems with power and flexibility by representing the world as a nested hierarchy of concepts, with each concept defined with respect to simpler concepts, and more abstract representations computed in terms of less abstract ones. It aims to solve the problems as done by the human brain [61, 185]. The human brain works with a connection of neurons, where it performs actions based on getting input signals

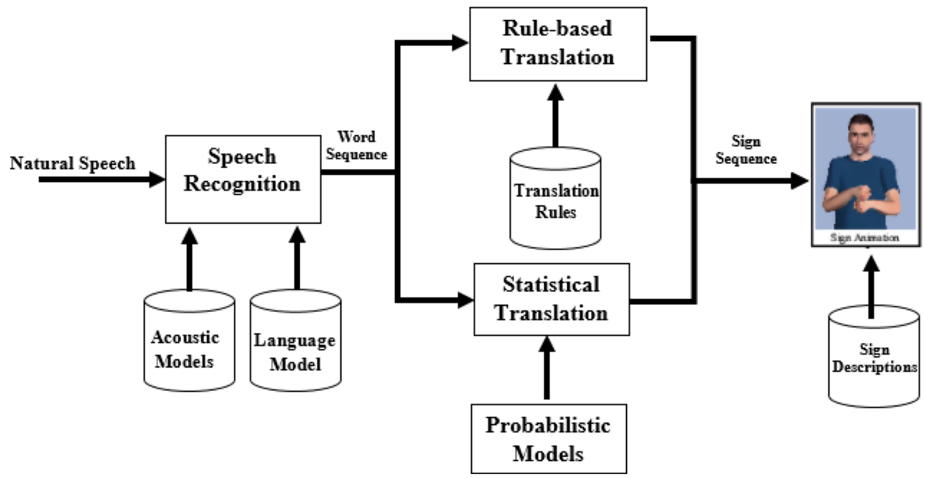


Fig. 17 Workflow for rule-based and statistical translation approaches [149]

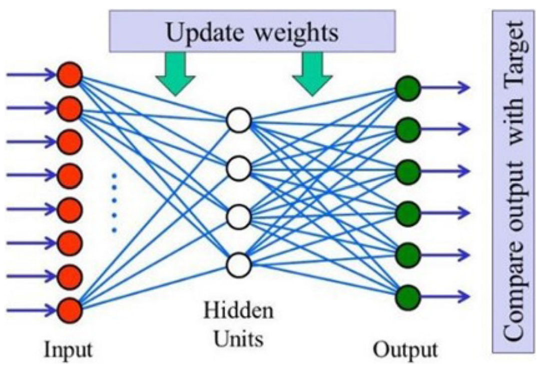


Fig. 18 Artificial neural network (ANN)

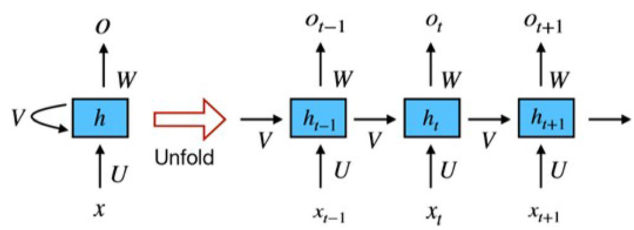


Fig. 19 Simple RNN network [163]

from the sensors. This approach is mimicked by modern technology, whereby it connects artificial neurons into a network using many layers, termed as artificial neural network (ANN).

Artificial neural network (ANN): ANN is a connection of neurons where some neurons are for input, and some are for output, while there can be one or more (hidden) layers that connect and process the input to translate it into output, as shown in Fig. 18. Each neuron receives an input, combines the input value with its internal state, and processes it using an activation function, and produces output using an output function. The output is then passed onto the neurons connected with it. Each connection has an associated weight which represents its relative importance in the network. One neuron aggregates all the inputs received from previous layer as a weighted sum and adds a bias to it to make and process its input. The network continuously learns to improve the handling of the task at hand, and the learning is mainly realized by adjusting the weights associated with each neuron so as to increase the accuracy of the neural network [193]. The ANNs work iteratively and keep on adjusting the weights by learning after each iteration. The learning rate governs the total number of iterations, where a higher learning rate seems to reduce the number of iterations, but it may lower the accuracy as well, while a lower learning rate takes more time, but potentially results into better accuracy. The weights of the network are generally adjusted using backpropagation. Technically, the backpropagation computes the gradient or derivative of the cost function associated with a given state with respect to the weights.

Deep learning approaches have been used for language translation problems [165]. Recurrent networks-based approaches including conventional recurrent neural network (RNN) [34, 39, 163], long short-term memory (LSTM) [184], and gated recurrent unit (GRU) [41] are the widely used approaches used for translation systems. Simple RNN, LSTM, and GRU models are presented in Figs. 19, 20, and 21, respectively.

Deep learning approaches have been used for language translation problems, where LSTM and RNN are two widely used approaches used for translation systems [165]. The reason for using these translation systems is that they tend to capture the context of the sentences using their memory structures and thus perform better than conventional statistical approaches, which struggle to incorporate semantic information in the absence of any suitable memory structure.

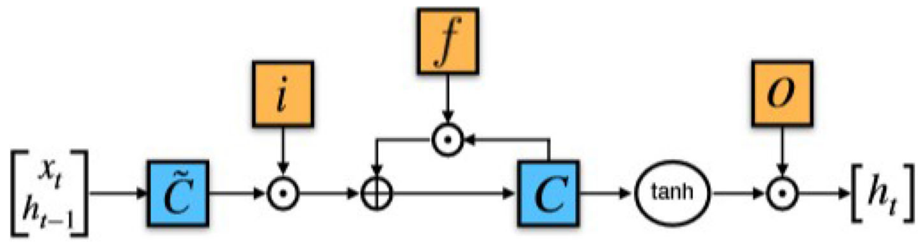


Fig. 20 Long short-term memory (LSTM) [184]

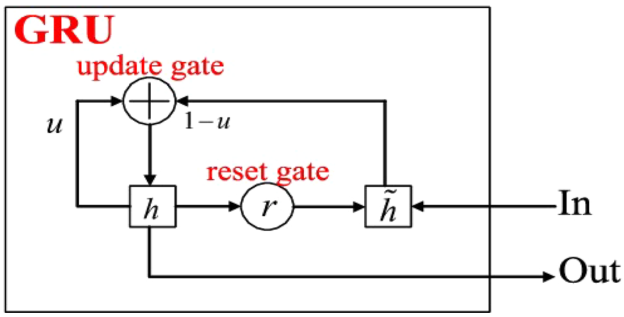


Fig. 21 Gated recurrent unit (GRU) [41]

5.3.1 Neural machine translation systems for natural to sign language translation

It is pertinent to mention that the deep learning approaches work well only when they are provided with a significant and reasonably inclusive training data set. There exist several machine translation approaches for conventional machine translation systems that convert data from one natural language to another natural language, because a large number of data sets are available. However, no sufficient data sets exist for the translation of natural language to sign language [31]. One of the early neural machine translation models was proposed by [126]. Later on, [33] used bidirectional encoder representations from transformers (BERT) for ASL, while [121] used transfer learning for translation of English into BSL. Another research that involves neural machine translation for DGS has been accomplished in [42] that involves a data set of weather news gestures. This work involves gesture recognition and then translates the identified gestures into German language using deep learning. The biggest challenge in the application of neural machine translation approaches for natural to sign language is absence of sizeable data sets for this purpose [31]. Another recent work by Stoll et al. [161] involves generative adversarial networks (GAN) for sign language translation.

5.4 Data sets for natural to sign language translation

A list of some prominent data sources used for natural to sign language translation for different sign languages is presented in Table 7. It can be observed that the number of sentences in these data sets is very small. Therefore, the existing deep learning-based translation models that convert natural language to equivalent sign language use generative models like BERT due to lack of available data sets. A comparison of speech data sets with the sign language data sets is presented in Table 8. The data in Table 8 show that sign language translation data sets are order of magnitude smaller than speech related data sets. Therefore, creation of a sizeable data set is one of the biggest challenges that needs to be addresses to design and test useful neural translation models.

5.5 Evaluation measures for natural to sign language translation

Though the eventual output of this translation should be presented in the form of sign language gesture, those gestures are actually rendered after getting this textual translation. Therefore, this intermediate textual representation is necessary for translation of natural language into equivalent sign language gestures. Word error rate (WER), translation error rate (TER), Bilingual Evaluation Understudy Score (BLEU), and Metric for Evaluation of

Table 7 Corpora for natural to sign language [135]

Corpus	Language pairs	Sentences
Othman and Jemni	English-ASL	431
Bonham	English-ASL	11,000
Cate and Hussian	English-ASL	600
–	Italian–Italian SL	585
RWTH-phoenix	German–German SL	2486
–	Chinese–Taiwanese SL	1983
–	Catalan–Catalan SL	199
ATIS	English-iris SL	595
ATIS	German–German SL	595
ATIS	English south African SL	595
ID	Spanish–Spanish SL	2000
DL	Spanish–Spanish SL	2000
eSign-3D	Spanish–Spanish SL	5000

Table 8 Comparison of speech and sign language corpus sizes [31]

Tasks	Natural to sign language translation	Sign language recognition, generation, and translation	Speech recognition, generation, translation
Typical articulated corpus size	< 5000 sentences	< 100,000 signs	
Typical annotated corpus size	< 5000 sentences	< 100,000 signs	5 million words
Typical corpus vocabulary size	< 500	1500 signs	1 billion words
What is being modeled	Sign language sentences	1500 whole signs	300,000 words
Typical corpus numbers of speakers	N/A	10	1000

Translation with Explicit Ordering (METEOR) are the widely used evaluation measures for natural language translation problems. The computational formulae and brief description of each of them are given below. BLEU is most widely used measure in the literature reviewed for natural to sign language translation problems. These measures help evaluating the translation of natural language text into equivalent sign language text.

Word Error Rate: WER is based on Levenshtein distance that works at the word level rather than at phoneme level. Equation 1 presents the formula for the computation of WER.

$$\mathrm{WER} = (S + D + I)/N = (S + D + I)/(S + D + C) \tag{1}$$

where S is the number of substitutions, D is the number of deletions, I is the number of insertions, C is the number of correct words, while N is the number of words in the reference, i.e., $N = S + D + C$.

Translation Error Rate: It is a measure that estimates the amount of postediting required for a machine translation output. The formula for calculating TER has been presented in Eq. 2.

$$\mathrm{TER} = \text{No. of edits/average no. of reference words} \tag{2}$$

Bilingual Evaluation Understudy Score (BLEU): It is a widely used measure for the evaluation of machine translation tasks. It evaluates the translated output segment, preferably a sentence, with available possible good quality reference translations. Subsequently, an average score is computed for the output of the whole corpus. In this process, intelligibility or grammatical correctness is not valued. It ranges from 0 to 1, where 1 means perfect match.

Metric for Evaluation of Translation with Explicit Ordering: METEOR is an improved form of BLEU where it evaluates the translation output by incorporating stemming and synonyms and thus results in relatively better correlation with human judgment scores as compared to BLEU.

5.6 Comparative analysis and limitation of natural to sign language translation

A comparative analysis of all the existing approaches and their efficiency is presented in Table 9.

Among the semantic translation models, direct machine translation approaches are not widely used because of their inability to preserve the semantics of source sentences during translation and hence can only be used for restricted domains, while the transfer-based machine translation is better choice, as it is more generalizable and extendable since it makes use of a set of linking (transfer) rules. Furthermore, unlike the direct translation approach it accommodates mappings and operations, such as swapping the object and the subject and elimination of articles due to which they can handle languages with different sentence structures and varying word orders. The flexibility and extensibility of this approach lie in the fact that we may add more rules to the system seamlessly or change any existing rules to refine them and improve the translation accuracy.

The statistical model-based approach is highly dependent on the sample data and favors the sentences and words which are frequently appearing in the data set. Table 9 shows that with a small data set even the statistical approaches are unable to exhibit promising results, rather they show less accurate translations as compared to the semantic approaches.

The data in Table 9 show that very few neural translation approaches have been proposed for natural to sign language translation. Among these, the approaches that use deep learning models for natural to sign language

Table 9 A comparison of natural to sign language translation systems

Project	Language	Approach	Domain/data set	Output type	Non-manual feature	Accuracy
Type of machine translation						
Semantic machine translation approaches						
ViSiCAST/TESSA [26, 54]	Speech to BSL	Direct translation	Post office encounters	Pre-recorded videos	Yes (as it used pre-recorded videos)	61% complete phrase and 81% for sign units
TEAM [190]	ASL	Synchronous Tree Adjoining Grammar (STAG)	Syntactic transfer based on rules	Animated character	No	Not available
ZARDOZ [170]	ASL	Transliteration	Some basic sentences	Spatial dependency	N/A	Not available
ASL Workbench [46]	ASL	Lexical Functional Grammar	Rule based	Animated character	No	Not available
eSign [3]	BSL and DGS	Rule based	Government official websites	Virtual avatar	To an extent	Not provided
INGIT [172]	ISL	Grammar	Railway ticketing	Virtual human	To an extent	In house testing
SASL (follows TEAM) [81]	South African SL	STAG	Police/Hospital	Avatar	No	Not provided
Statistical approaches						
eSign-3D [147]	LSE	Statistical and rule based	5000 sentences dialogues (Driving license)	3D Avatar	To an extent	BLEU Score rule based: 0.57 Statistical: 0.49
Othman and Jemni [125]	ASL	Statistical	431 sentences	Text to text	N/A	Not provided
Bonham [28]	ASL	Statistical	~ 11,000 sentences Not provided	Text to text	N/A	BLEU 0.17
Krnoul [101]	Czech SL	Statistical		Avatar	To an extent	Perceptual test
Neural Machine Translation (NMT)						
ASL-Eng (Cate and Hussain) [33]	ASL	Deep learning BERT (generative model)	~ 600 sentences from lifeprint.com	Text to text	N/A	BLEU 0.18
DGS-German [73]	DGS	NMT	Weather forecast videos	Video to text	No	BLEU 0.14
ASL-Eng [83]	ASL	NMT	ASLG-PC12	Text to text	N/A	BLEU 0.17
Mocialov [120]	BSL	Transfer learning with LSTM	Penn Tree Bank	Text to text	N/A	121.46 perplexity

translation are based on transfer learning or are using generative models; i.e., they are customizing the already trained models developed for some other similar problems, like a model used for translating one natural language to another natural language. Furthermore, like statistical models these approaches are also data hungry and require big data sets for better results. It is evident from the comparison shown in Table 9 that neural machine translations approaches also exhibit very low accuracy as shown by their BLEU scores.

5.6.1 Major limitations

The thorough analysis presented above reveals the following major limitations in the natural to machine translation systems.

- Most of the semantic approaches are domain specific and cannot be generalized for all domains.
- No sizeable and generic data sets are available which may be helpful in proposing generic and more accurate

statistical or neural machine translation models for natural to sign language translation.
- The accuracy of all types of translation systems is very low and need to be improved.

6 RQ3: gesture recognition

Gesture recognition is generic research problem [178], while in the context of sign languages it involves the identification of a sign language gesture with the help of available datasets using a computer system [118, 176, 189]. These approaches can be categorized into two broader types, namely hardware-based approaches and software-based approaches, where the hardware-based approaches are based on Glove, Kinect, or any other sensor-based device. On the other hand, the software-based approaches are based on computer vision, image processing, and use probabilistic models or machine/deep learning models for gesture recognition.

Gestures of a sign language represent the written natural languages with almost similar granularity. For instance, American Sign Language (ASL) alphabet refers to 26 finger-spelled letters that includes 24 static gestures and 2 dynamic gestures which are “J” and “Z” [133]. Remember the static gesture does not include any movement in the sign, whereas dynamic gesture contains the movement of hands, while the next granularity level is that of words or phrases which have different levels of complexities based on the types of gestures, i.e., single-handed static or dynamic gestures to double-handed static or dynamic gestures.

6.1 Hardware-based approaches

Hardware-based approaches for gesture recognition involve some specific wearable hardware, sensor-based systems, or camera-based systems that provide some additional information than the tradition video cameras. We divide these hardware-based approaches into the following three categories:

- Glove-based systems
- Kinect-based systems—sensor-based systems

6.1.1 Glove-based systems

Gloves are made up of some sensors that are used for acquiring data, electronic modules, and power supply, which collectively help in data acquisition and processing [50]. As the name suggests these gloves are worn on hands while performing gestures. Essentially, it is an old technology that was invented in 1970s and has been used for several different purposes including robotics, manufacturing, healthcare, and also in sign language recognition, whereby Gary Grimes developed first data glove in 1983 for the purpose of gesture recognition [70], which was followed by a long streak of different variants of data gloves to support the deaf community. These gloves include digital entry data glove [70], CyberGlove [1], Accele Glove, LightGlove, power glove, chording glove, and smart glove.

These gloves are different from one another based on sensor types and positioning. Furthermore, different recognition techniques have been used with different gloves. Similarly, different sensors are useful for different types of gestures and offer varying accuracy for static and dynamic gestures. Table 10 enlists different glove-based gesture recognition approaches for different number of gestures for various sign languages. The table also shows the number of gestures recognized and the accuracy of the approach.

Most of the early work in gesture recognition has been accomplished for ASL. A lot of research has been conducted in the field of recognition of ASL. Smart glove was used to recognize 10 alphabets of ASL by Praveen et al. [137]. Subsequently, 24 static alphabets of ASL were identified by Mehdi et al. [119] using the sensor glove technology. Das et al. [47] proposed a smart glove-based approach for the recognition of all 26 alphabets of ASL. Data glove with some additional sensors was used by Parvini et al. [131], while Lee et al. [104] designed fusion gloves to recognize 26 alphabets of ASL. Ambar et al. [21] designed flex sensors to recognize 6 static gestures-based words of ASL.

On the other hand, some of the researchers recognized words which are spelled using finger spellings. As a matter of fact, the proper nouns including the names of a person, place, or the words with no designated sign are finger spelled by the deaf persons, where the signer performs the sign of each alphabet in the word. The finger spelling was recognized by Haydar et al. [74] while using flex sensors; however, it was not tested on word level. Oz et al. [129] developed a system that uses data glove to recognize the names which are finger-spelled.

Though significant work has been done on ASL, glove-based approaches have been developed and test for many other languages as well. Some elementary work on the gesture recognition of Spanish, Indian and other sign languages has been conducted by different researchers. Angulo et al. [174] used data glove with ten Flex sensors, while accelerometer glove was used by Gonzalez et al. [148] for the recognition of 26 Spanish Sign Language alphabets. Krunal et al. [156] recognized the ISL alphabets

Table 10 Glove-based approaches

Approach	Language	Word	Total no of gesture (size)	Accuracy (%)
[137]	ASL	10 alphabets	10	–
[119]	ASL	24 alphabets	24	88
[131]	ASL	24 alphabets	456	Leave-one-out: 73.39 Keep-one-in: 43.49
[104]	ASL	26 alphabets	6240	98.2
[129]	ASL	26 alphabets	–	96
[47]	ASL	26 alphabets	3	86.67
[74]	ASL	26 alphabets	26	–
[21]	ASL	9 commonly used words	–	Average sign detection time is 0.74 s
[174]	LSE	26 alphabets	250	Precision: 97.39
[148]	LSE	26 alphabets	1378	97
[156]	ISL	26 alphabets	26	–
[109]	ISL	6 words	32	99
[89]	PSL	10 words	300	90
[122]	ArSL	100 signs	2000	99.6
[110]	CSL	10 numbers	10	89.59
[85]	CSL	19 signs including numbers, alphabets, and 2 words	50	91
[115]	CSL	220 words and 80 sentences	300	98.2
[153]	MSL	A-C, 1–3, and three words (total = 9)	90	89
[166]	MSL	25 signs	500	–

using the flex sensor glove, whereas data glove was used by Lokhande et al. [109] for the recognition of 6 ISL words.

Mohandes [122] recognized 100 Arabic Sign Language words by using two cybergloves. In Pakistani Sign Language, Kanwal et al. [89] used flex sensor, accelerometer, and arduino for the construction of glove. The correctly classified sign was displayed on LCD attached with glove.

Lu et al. used data glove [110] to recognize numbers in CSL. Jingqiu and Ting used data glove [85] to recognize alphabets from A to G and numbers from 0 to 9, and they also recognized two words. Ma et al. [115] did a considerable work in gesture recognition for CSL, whereby they used 2 cyber-gloves and two 3space position trackers to recognize 220 words from CSL and evaluated it on 80 sentences of CSL.

Some work has also been accomplished for Malaysian Sign Language, where researchers have used glove-based approaches to recognize alphabets and words. Shukor et al. [153] used data glove to recognize 3 alphabets, 3 numbers, and 3 words from MSL, while Swee et al. [166] recognized 25 words from MSL by using two data gloves and accelerometers.

The glove-based gesture recognition methods have been used to recognize the gestures for different granularity levels for many different sign languages of the world. Figure 22 shows some sample gloves that can capture finger bends. These approaches work well, but there are many limitations associated with these approaches that include portability and cost. The approaches exhibit poor robustness. Lastly, these approaches disregard the non-manual features altogether, as they only focus on manual features as they rely on the data obtained from the gloves worn on hands.

6.1.2 Kinect-based approaches

Microsoft © has provided a hardware device that senses the motion of human body and plots different body points in a 3D space [112]. It was initially released in November 2010, where the initial focus was on playing games and entertainment, while an upgraded version and associated software development kit (SDK) were released in 2014. This SDK along with third-party software toolkit, e.g., OpenNI, helps capturing different streams in the form of colored images, and 3-dimensional depth and skeletal frames. Figure 23 shows how a gesture is captured by MS Kinect.

This device has also been used for motion sensing and body movements [112, 194]. Microsoft Kinect is a low-cost motion sensing device and facilitates user to interact with a computer through gestures and voice commands. Particularly, SDK version 2 has also been used in a variety of

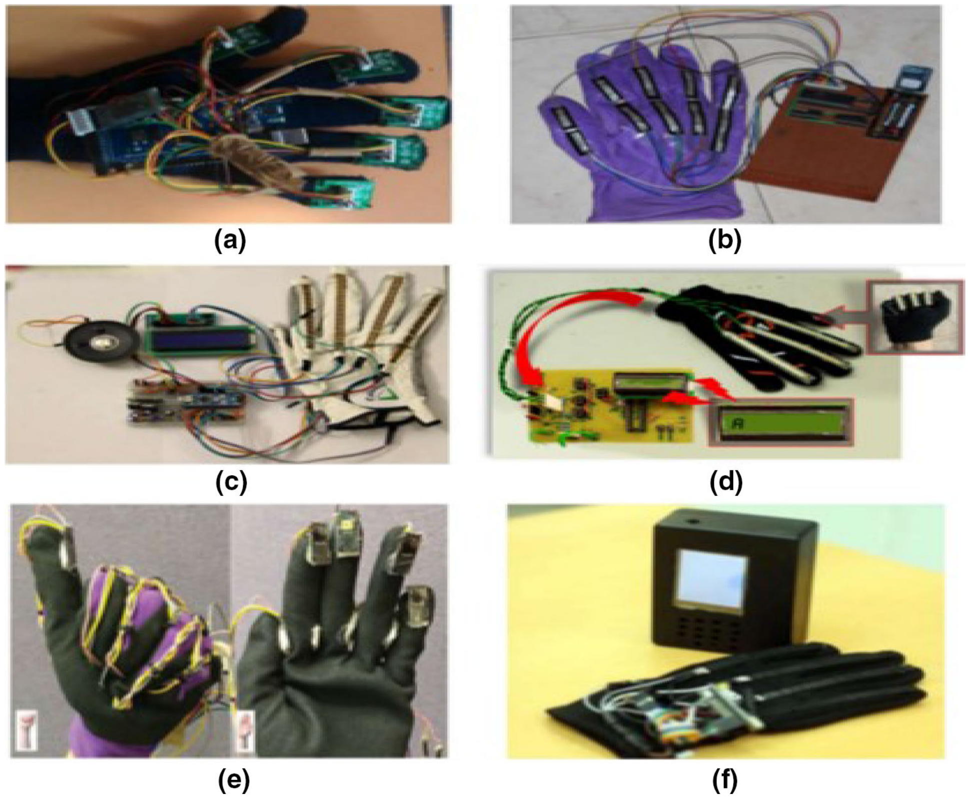


Fig. 22 Finger bend detection glove used in the literature. **a** The glove consists of five 3-axis ACCs and **b** the ten custom made flex sensors; **c** the translator system consists of five flex sensors mounted on glove, LCD, and speaker; **d** the glove consists of three flex sensors used to detect few gestures; **e** the glove is equipped with five contact (pressure) sensors; **f** the wireless translator system is embedded with the five flex sensors and LCD [17]

areas including education, robotics, performing arts, etc., while involving machine learning.

Kinect-based approaches have also been used to develop system that help in gesture recognition for different sign languages [103, 112, 162] on varying level of granularity of gestures including alphabets, words, and phrases. These approaches along with sign language, the number of words, the total number of gesture data set, and accuracy are mentioned in Table 11.

Many different efforts have been made on the recognition of different sign languages including ASL, DGS, PSL, CSL, LSB, and ISL as shown in Table 3. The process of gesture recognition using Kinect involves features extraction from the data streams produced by Kinect, followed by applying appropriate classification technique to correctly identify the gesture, while different data sets have also been developed to evaluate the approaches for different sign languages. In almost all the cases, accuracy has been

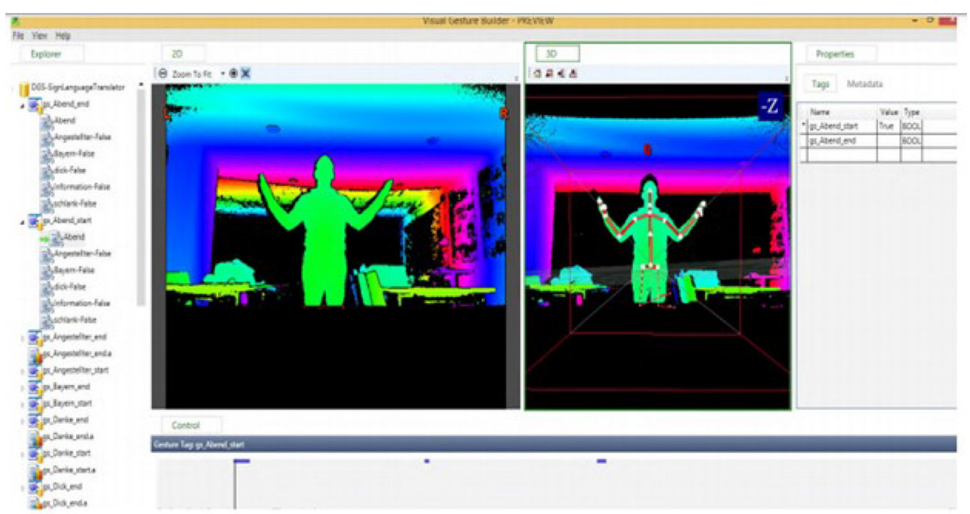


Fig. 23 Gesture being captured by MS Kinect

Table 11 Kinect-based approaches

Approach	Language	Word	Total no of gesture (size)	Accuracy (%)
[109]	ASL	24 alphabets	72,000 depth images	Half-half 90% Leave one out 70%
[140]	ASL	26 alphabets	48,000	Vector (mean precision 75%), appearance (mean precision 73%) and depth (mean precision 69%)
[186]	ASL	24 words	240	90.4%
[188]	ASL	60 phrases	1000	Seated position 51.5% Standing position 76.12%
[103]	DGS	25 signs	40 to 60 against each sign	97%
[66]	ISL	10 words	6 samples (2 best cases, 2 average cases, 2 worst cases)	Best case 100% Average case 40% Worst case 25%
[72]	PSL	20 Words	80	91%
[16]	CSL	10 numbers		Max accuracy 92.313
[162]	CSL	73 words	1971 phrases	86.8%
[35]	CSL	239 words	1195	Rank 1 83.51% rank-5 96.32%
[19]	LSB	34 signs	–	Avg 80

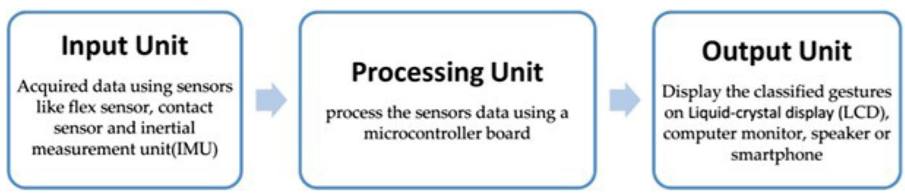


Fig. 24 Steps involved in sensor-based gesture recognition approaches

used as the evaluation measure to compute and compare the approaches. The accuracy obtained by different approaches is reported in Table 11. The approaches have used the features either in the raw form, i.e., skeletal joint information, or 3D depth information from hand motions, while some have used other filtering techniques like Gabor filter, axis of least inertia, etc.

A variety of classification approaches have been used to classify the gestures, and many of them have used random forest, hidden Markov models, hierarchical conditional random field, dynamic time wrapping, discriminative exemplar coding, 3D trajectory encoding algorithm, and support vector machine, and some deep learning approaches.

Lastly, the researchers have created different data sets for gesture recognition using Kinect for many different sign languages. Some of them are just limited to static alphabets, while other include all alphabets and the more credible ones involve words- and phrases-based gestures, which in turn can be used for sentence-level gesture recognition. The statistics of different data sets used in different studied approaches are presented in Table 11.

6.1.3 Sensor-based systems

Apart from glove and Microsoft Kinect, the researchers have used a few other sensors as well for human action recognition. These sensors include leap motion sensor, electromyography (EMG), and accelerometer, and depth sensors. The sensor-based approaches are generally composed of three phases: (a) input; (b) processing; and (c) output, as shown in Fig. 24.

The leap motion sensor [12, 25] has been used to figure out the hand movement. It is small low-cost device with the capability to capture and distinguish fingers, joints, and movement. It is used by a huge number of developers worldwide. It is pertinent to mention that in contrast with Kinect, it only tracks hand movements. Electromyograph [10] is used to capture skeletal muscle activity. The signals are used to detect movement of an object. Accelerometer [5] is an electromechanical device, used to measure acceleration forces, which by definition is the rate with which the velocity of an object changes over time.

Table 12 Sensor-based approaches

Approach	Language	Sensor	Classification Approach	Word	Total no of gesture (size)	Accuracy (%)
[168]	ASL	Leap motion	K-NN	24 alphabets	120	Leave-one-out: 80.1 half-half experiments: 99.7
[40]	ASL	Leap motion	K-NN	26 alphabets	7900 observations	k-nearest neighbor: 72.78 support vector machine: 79.83
[183]	ASL	EMG sensor		40 commonly used words	1000	95.94
[93]	DGS	EMG sensor	K-NN	7 words	490	99.82
[102]	ISL	Leap motion + kinect	HMM	25 words	2000	90.80
[56]	ArSL	Leap motion	MLP	50 signs		88
[191]	CSL	Accelerometer + EMG	HMM	72 words	8640 words	Words 93.1
					800 sentences	Sentences 72.5
[105]	CSL	Accelerometer + EMG		121 sub words	600	Sub words 96.5
				200 sentences		Sentences 86.7
[136]	Auslan	Leap Motion		26 alphabets	–	–

These sensor-based approaches use different classification approaches including SVM, multilayer perceptron, K-nearest neighbor (K-NN), hidden Markov model (HMM). Table 12 presents a categorization of these sensor-based approaches based on the language, classification approach, type of sensor used, granularity and number of gestures, total number of gestures, and accuracy.

Discussion: The hardware-based gesture recognition techniques can be classified into three main categories: glove-based, Kinect-based, and other sensors-based approaches. The glove-based approaches require the users to wear the gloves in their hands, where these gloves help capturing the data related to shape, location, orientation of hand for the static gestures mainly comprising of alphabets, while some of them handle dynamic gestures with single or both hands as well which also capture the motion data.

An extensive literature review has been conducted to figure out the methodologies, tools, and devices used to recognize gestures of sign languages. The initial data collection quickly revealed that most significant and sophisticated work in this regard has been conducted on ASL, while recently a noticeable amount of work has been conducted in Asian sign languages including CSL, ArSL, ISL, PSL, MSL, etc. It can be clearly observed that three different types of hardware-based approaches have been used to recognize gestures of a sign language. Firstly, glove-based approaches have been used, where people have devised gloves that are worn by the signer and these gloves provide information related to the shape, orientation, location, and movement of a person's hands, while the sensor-based approaches also involve different sensory material that is worn on wrist or hands, and this material provides certain useful input data to train the automated system to learn gestures and thereafter recognize it. Lastly, Microsoft Kinect-based solutions also exist, which captures RGB image along with skeletal information of a person's joints. A reasonable accuracy has been achieved with this device as well, but it does not provide the information of all hand joints and their positions which hurdles its accuracy to determine similar gestures. The summary tables clearly show that the research on gesture recognition in ASL involves bigger data sets and shows reasonable accuracy, while in Asian languages the researchers have built small data sets and have obtained high accuracy values. There is a need to build bigger data sets for the Asian languages as well. Similarly, it can be observed that glove-based approaches exhibit better results than Kinect- and sensor-based approaches. The reason is that glove covers most parts of the hands and transmits necessary information to figure out the hand shape in a better way, thus resulting in a better accuracy for similar gestures.

Another important dimension is the granularity of gesture recognition systems. The granularity in this case depends upon the letters, words, phrases, and sentences. Most of the work done in gesture recognition revolves around the recognition of letters and numbers. A very small amount of work is done on the recognition of words, while translation of recognized words into sentences is yet another difficult challenge that still needs to be accomplished even in ASL. Lastly, current gesture recognition approaches totally disregard non-manual features.

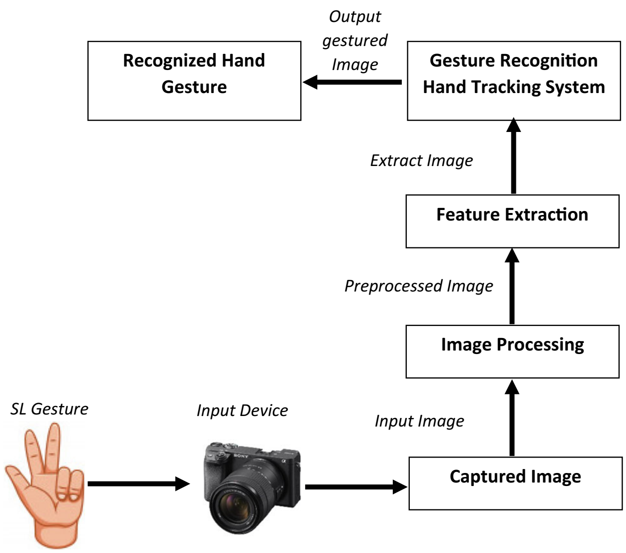


Fig. 25 Steps involved in software-based gesture recognition approaches [17]

6.2 Software-based approaches

There are a variety of approaches that do not use any hardware to recognize the gestures of a sign language. These approaches mainly fall in the area of computer vision and rely on digital image and video processing while using machine learning and deep learning approaches for gesture recognition. These approaches use images or video data set for training and testing purposes. The software-based approaches can be further categorized into probabilistic approaches and machine/deep learning-based approaches. A high-level workflow for software-based approaches, which are also considered to be computer vision-based approaches, for gesture recognition is presented in Fig. 25.

Machine learning-based approaches rely on one or more statistical models, such as the hidden Markov model (HMM), support vector machine (SVM), and artificial neural networks (ANNs) to recognize the sign gesture (Fig. 26).

6.2.1 Probabilistic and machine learning approaches

Hidden Markov model (HMM) [142] is a statistical Markov model in which the system being modeled is assumed to be a Markov process with un-observable states. The mathematics behind the HMM were developed by L. E. Baum and coworkers. HMMs are used in speech recognition system, pattern recognition, areas of artificial intelligence (AI), and in computational molecular biology. In recent times, HMMs are also been used in computer vision such as image sequence modeling, object tracking, and recognition.

To recognize sign language, several HMM- and AI-based approaches have been developed. These approaches take input in the form of image or video and extract hand features based on image processing techniques, and based on these features, sign is recognized. The developed approaches along with sign language, the number of words, the total number of gesture data set and accuracy are mentioned in Table 13.

Starner et al. [159] recognized ASL sentences consisting of 40 words lexicon by using HMM classifier. For the acquisition of sign images, hat mounted wearable camera and front desk camera were used. Hat camera is used to capture complete finger orientation from the top, and desk camera is used to capture full sign view. By using hat mounted camera, they achieved 98% accuracy and desk

Fig. 26 Google AI is paving way for mobile applications for sign language gesture recognition

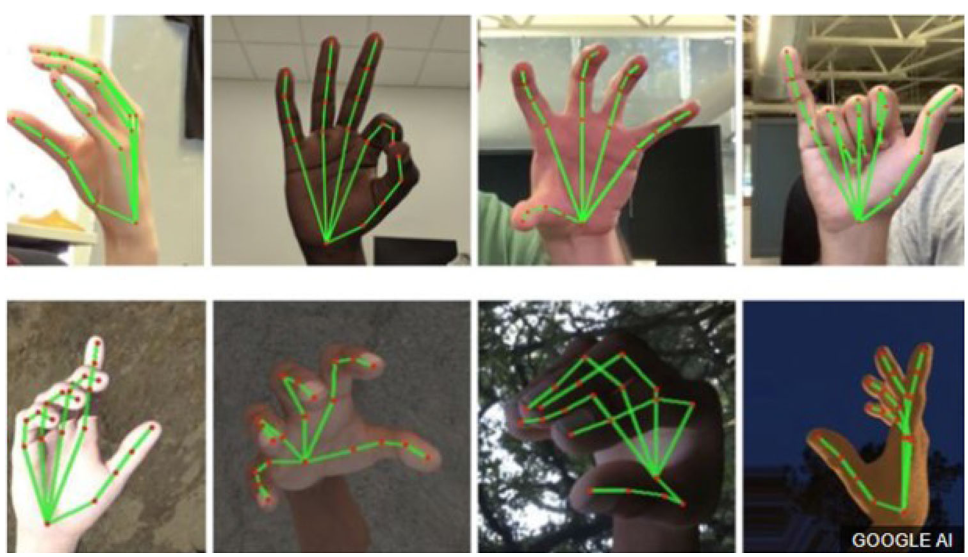


Table 13 Probabilistic and machine learning-based approaches

Approach	Language	Word	Total no of gesture (size)	Accuracy (%)
[158]	ASL	40 words	494	99.2
[159]	ASL	40 words	500	98
[108]	BSL	100 words	1000	98.9
[145]	ISL	17 words	2846	90
[132]	LSB	15 words	180	91.1
[27]	LSB	40 words	9600	96.77
[114]	ArSL	23 words	3450	98.8
[82]	ArSL	30 words	450	97

camera achieved 92% accuracy. Starner et al. [158] applied HMM-based system for classifying sentence of American Sign Language and achieved 99.2% word accuracy. Liwicki and Everingham [108] recognize 100 words in BSL by tackling finger-spelling approach. Source of input in this work is webcam, and signer performs sign by using both hands. HMM model is used for the recognition of sign.

Rao and Kishore [145] use smart phones for the recognition of signs of ISL. For this purpose, selfie camera was used. From the captured images, head and hand features were extracted. Minimum distance and artificial neural network classifiers are used to classify signs from the extracted features.

Bastos et al. [27] recognize static signs of Brazilian Sign Language by using artificial neural network. Pizzolato et al. [132] employed finger-spelling approach for the recognition of Brazilian Sign Language words. They take video input form camera and then classify using artificial neural network.

Sidig et al. [114] recognized 23 words of Arabic Sign Language consisting of both hands. They compare three transformation techniques (viz. Fourier, Hartley, and log-Gabor transforms) and three classifiers results (KNN, MLP, and SVM) results. According to the study, Hartley transform and SVM give best results. Ibrahim et al. [82] used Euclidean distance algorithm for the recognition of 30 Arabic sign language words. 450 videos with white background are captured for the training and testing purpose, whereas, recently Brock et al. [32] managed to recognize non-manual content in continuous Japanese sign language video streams.

The accurate tracking of individual hand and finger movements has been notoriously difficult to master. However, a recent project from Google's AI labs reveals that they can infer 21 points on a person's hands based on palm's position. This can help in developing applications for gesture recognition. Figure 26 shows the screenshot of Google AI system for hand gesture recognition.

A significant amount of research work has been done by using software-based approaches for sign language gesture recognition. These approaches either involve probabilistic methods which are mostly based on hidden Markov models, and some are based on machine learning approaches including SVM, KNN, and neural networks. These approaches have shown very good results in gesture recognition which is up to 98% accuracy for 100 word gestures data set. On the other hand, deep learning-based approaches have started gaining more popularity for showing better results and scalability. Most of these approaches use convolutional neural networks, which have shown promising results in various computer vision applications, in general. The best approach shows up to 92% accuracy for 900 word gestures. Popular public sign language video corpora used for sign language recognition are presented in Table 14.

6.2.2 Deep learning-based approaches

Many deep learning-based approaches have also been proposed for gesture recognition for sign languages. A summary of these approaches is presented in Table 14. The table shows that some fundamental work involved simple gesture recognition that involves alphabet and digits recognition for different languages [84] ASL, [154] ISL, and [52] ArSL, while some recent work manages to achieve continuous gesture recognition using state-of-the-art data sets [97, 139] for DGS and CSL (Table 15).

Jalal et al. [83] proposed framework works on ASL static letters. By using convolutional layer and spatial adaptive pooling with capsule network routing, sign was recognized. The 27455 images from Kaggle' American Sign Language Letter database were used for training and testing purpose. Kim et al. [94] used to recognize ASL words by using finger-spelling approach. Their major contribution in this regard is to extract appropriate frame from the video sequences containing fast and short interval alphabets signs. To make it easier to extract sign, signers wore specific color clothes with green background while

Table 14 Popular public sign language video corpora used for sign language recognition

Data set	Vocabulary	Signers	Signer independent	Video	Continuous	Real-life
Purdue RVL-SLLL ASL [117]	104	14	No	2567	Yes	No
RWTH Boston [187]	104	3	No	201	Yes	No
Video-based CSL [77]	178	50	No	25,000	Yes	No
Signum [174]	465	25	Yes	15,075	Yes	No
MS-ASL [86]	1000	222	Yes	25,513	No	Yes
RWTH-Phoenix [62]	1081	9	No	6841	Yes	Yes
RWTH-Phoenix SI5 [98]	1081	9	Yes	4667	Yes	Yes
Devisign [36]	2000	8	No	24,000	No	No

recording the sign. Convolutional neural networks were used for the recognition of sign. With the neural network adaption, they recognized signs containing multiple signer. Jin et al. [84] created a mobile application for the recognition of 16 static alphabets of ASL. In the developed application, canny edge detection algorithm was used to extract sign feature. To classify sign from the extracted features, SVM is used. Singha et al. [154] recognized ISL alphabets consisting of both hands. Euclidean distance classifier is used to classify hand features, extracted from video frames. Tripathi et al. [170] proposed a continuous ISL gesture recognition system. In this continuous gesture recognition system, 10 sentences videos containing 2–3 gestures are captured. From the captured videos, right frames are extracted using key-frame method. These frames are then used for the classification purpose. A comparative analysis between different classifier (Euclidean, Mahalanobis, City block, and correlation distance) is presented in the study. According to experimental results, Euclidean and correlation distance gives better results. In Athira et al. [24], the authors have designed and developed a system for the recognition of finger-spelling words for ISL from live videos. The data set used in the study consists of static, dynamic single-handed gestures and double-handed static gestures. All these gestures, images, and videos are captured by using static background. SVM classifier is used for the recognition of sign. In El-Bendary et al. [52], the authors created a desktop-based application for the recognition of ArSL alphabets. This system is named as Arabic Sign Language Alphabets Translator (ArSLAT). This system takes input in the form of video and after pre-processing [143], feature extraction [144], signs are classified. Minimum distance (MDC) and multilayer perceptron (MLP) classifier is used to recognized sign. However, most recently some significant work has been done by [45, 139], and [97] which employee state-of-the-art data sets for continuous sign language recognition and have shown significant improvement in their results, where [45, 139] employ iterative training and iterative alignment, respectively, while using deep learning for continuous sign language recognition, whereas [76, 97, 120, 177] use deep learning models for gesture recognition of sign languages.

7 RQ4: avatar technology

An avatar is an image that represents one party in a collaborating argument. In broader sense, the avatar may represent an actual human being. The avatar almost always operates as an agent of any application and generally simulates human activity [95, 175]. An avatar has a built-in element of interactivity. It responds to the users' requests and needs, it provides a clear, insightful, rapid link to an information database, and it does so in a manner that is easy to understand.

An avatar drives as a communications interface connecting a user with the information the user desires. An avatar is neither a cartoon model nor a video demo. An avatar has a built in component of collaboration. It rejoins to the users' requirements and wishes, it delivers a vibrant, perceptive, swift association to an information database, and it organizes in a way that is easy to comprehend [18]. Avatars are pretty more significant where interactivity, learner engagement, and cultural features are central design considerations. Avatars are also exposed in the form of bots, chatbots, or infobots.

Technically, the pipelines that generate avatars usually involve a series of complex steps. The animations for individual signs are usually rendered from a collection of individual signs. The motion plan for a sign can be produced in many ways, e.g., key-frame animations (e.g., [78]); symbolic encoding of subsign elements (e.g., [51]); or motion-capture recordings (e.g., [67, 151]), while the non-manual signals are fetched from complementary data sets (e.g., [79]) or can be synthesized from models (e.g., [87]). Subsequently, both manual and non-manual components are combined to generate a moving gesture. Apart

Table 15 Deep learning-based approaches

Approach	Language	Word	Total no of gesture (size)	Accuracy (%)
[45]	DGS, CSL	RWTH-Weather, CSL	6841	15% improvement on large data sets
[139]	DGS, CSL	RWTH-PHOENIX-Weather and CS	4667	0.98 BLEU, 0.37 WER
[97]	DGS, CSL	RWTH-PHOENIX-Weather and CS	6841	0.24 WER
[84]		16 alphabets	1600	97.13
[83]	ASL	24 alphabets	32427	99
[94]	ASL	900 words	3684	Single signer 92% multi signer 83%
[154]	ISL	24 alphabets	240	96
[24]	ISL	26 alphabets and 11 words	1600	91
[170]	ISL	10 sentences	100	Max 93
[52]	ArSL	30 alphabets	2000	Minimum distance classifier 91.3% multilayer perceptron classifier 83.7%
[73]	DGS	RWTH-PHOENIX Weather data set	3800	38.3% WER
[62]	DGS	SIGNUM: 465 RWTH-PHOENIX single signer:266 RWTH-PHOENIX Multi-signer: 1080	SIGNUM 14000 RWTH-Phoenix single 3800 RWTH-Phoenix Multi 71000	SIGNUM 7.4% WER RWTH-PHOENIX 2012 Single Signer 30% WER RWTH-PHOENIX 2014 Multi-signer 32.5% WER

from this, speed, timing, and sizing are also important aspects which are set by a human or may be controlled by rules [51] or can be predicted using a machine learning model [18, 111]. In the end, animation software is used to render the relevant components according to the settings.

A sign language avatar is a realistic portrayal of a character fit for depicting sign language through three-dimensional movement [20]. This is relatively new area as compared to the other sign language translation components. An avatar can fill various needs and is an essential part of any programmed interpretation framework that changes over a communicated in language into sign language. They are likewise utilized for training and for anonymization of online videos. The objective is to encourage better correspondence among hard of hearing and hearing communities and to make practically proportionate access to data, media, education, training, job opportunities, and social administrations [20]. Researchers have generated avatars for different sign languages [29, 53, 71, 179] and for different applications [22, 150, 152, 175].

The signing avatars started with two leading and influential European project ViSiCAST and eSIGN [3, 26, 91] which developed avatar technology, based on a sign writing notation HamNoSys, which is first converted into SiGML and subsequently into an avatar using a technology named Animgen. But this technology had limitations; firstly, it was not an open source technology, and secondly, it strongly depends upon HamNoSys that does not have any transparent mechanism to change animations. In order to address this limitation, another avatar Zebdee was developed by the LIMSI institute that helps generating avatars by parameterizable script [60]. However, this approach was hard to use and understand. Paula [181] and Greta [124] were other similar projects, but all of them could not be available for general use. EMBR [75] introduced the animation layer to de-couple the animation parameters and behavior specifications. It was further extended by [95] to enrich user experience, and non-manual features. The EMBR character in comparison with original human is presented in Fig. 27.

One major issue arises in the idea that an avatar simulates human activity. It means in an avatar must look like a human in appearance. Researchers have been performing a comparison of the avatars with human gestures [30, 37, 147]. These avatars are created using different animation tools including Maya3D, Blender, Zbrush, OpenGL, DirectX, Unity, and Unreal 4. Similarly, researchers have generated markup languages [29] for automating the avatar generation process in a dynamic and flexible manner.

Technology companies have also been developing avatar systems for the deaf community (Figs. 28, 29). For instance, Microsoft has created a multilingual avatar-based translation system that translates natural language to 8 different sign language avatars [2].

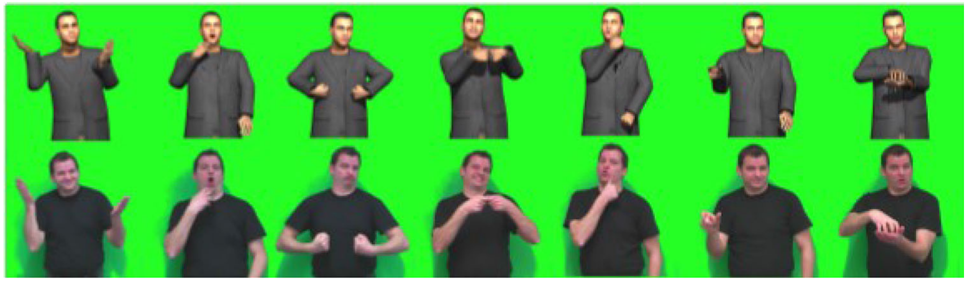


Fig. 27 EMBR character animation system [95]

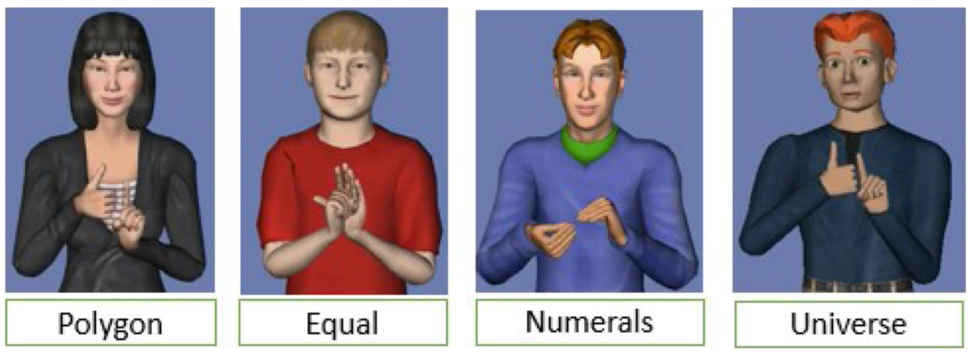


Fig. 28 Science and math dictionary by TERC

Fig. 29 Microsoft multilingual translator; spread the sign multilingual dictionary

The state of the art in avatar generation is not fully automated; all parts of current pipelines currently require human intervention to generate smooth, coherent signing avatars. Prior research has measured the quality of avatar animations via perceptual and comprehension studies with deaf participants [79], including methodological research [88] and shared resources for conducting evaluation studies [78].

8 RQ5: sign language translation applications

Several applications have been built by developers to learn and practice sign languages (Table 16). Similarly, others have made applications for sign language translation that involve natural to sign language translation, as well as the gesture recognition systems. Lastly, some domain-specific applications have also been developed for daily life activities including hospital, religious services, parenting, etc.

In this section, we present various different applications developed for all the aforementioned purposes. Most of the language learning applications focus ASL, while there exist applications to learn and practice other sign languages as well. Similarly, there are some domain-specific applications in different languages. Lastly, some translation applications have been presented as well.

Now, there is a need to integrate the gesture recognition systems with the language translation systems to build applications that help translating the gestures performed by a signer into equivalent natural language

Table 16 Various types of learning applications

Sr. #	Android (A)/IOS (I)/Both (B)	Paid (P)/ Free (F)/ Both (B)	Language	Application Name	Description	Best For
Language learning applications for ASL						
1	B	B	ASL	My Smart Hands Baby Sign Language Dictionary	Create by the makers of My Smart Hands which is an international program for helping the babies learn sign language. It presents a dictionary-based learning of sign language by involving more than 150 instructors that teach in more than 10 sign languages	ASL Learning for Kids
2	A	F	ASL	Sign Language: ASL Kids	The app has a simple interface based on taps and swipes that children should find easy and intuitive. Its library isn't huge, since it's focused on just teaching the basics, but it does that well	ASL Learning for Children
3	A	F	ASL	Sign Language for Beginners	Signing common letters, words, and numbers in ASL Practice and quizzes	ASL Learning for Beginners
4	B	F	ASL	Drops	Learn ASL basics by associating illustrations of the signs to their meanings. It also contains game-based tests	ASL learning and practice
5	B	F	ASL	Scripts	Helps learning ASL basics	ASL learning
6	B	B	ASL	The ASL app	This app contains more than 1000 videos that help learning ASL with fun and ease	ASL Learning
7	B	F	ASL	Signing Savvy Member App	An ASL dictionary comprising of few thousands gestures in the form of high resolution videos	A big multimedia dictionary for ASL
8	A	F	ASL	Sign ASL	This app is as barebones as they come. Type in a word or phrase, and it returns videos of people signing that word or phrase. There's only a tiny banner ad along the bottom, and each word usually has multiple video results so you can see multiple pronunciations of the same word	Videos of real people signing
9	B	B	ASL	Hands On ASL	It is an app that helps learning sign language alphabets using 3D models that helps viewing signs from all angles using finger spellings with customized skin colors	Learning sign language alphabets
10	I	P	ASL	ASL Coach	This app helps learning alphabets and numbers up to 9 in a coaching mode where the users can watch and repeat after the instructions	Learning alphabets and numbers
11	I	P	ASL	ASL Finger-spelling	It offers practice to learn finger spelling with varying length words and different gestures speeds. It also offers different modes to test one's language capabilities	Finger spelling learning and practice
12	I	P	ASL	Marlee Signs	It is an applications that has engaged an academy award winning deaf actress, Marlee Matlin, who teaches the fundamentals of ASL Lesson are customized for each individual so that they can learn at their pace, while keeping track of completed lessons	Learning through person
13	A	P	ASL	Memorize	It is one of the most famous language learning tools on Google Play It supports ASL as well	Language Learning tools that supports ASL as well
14	A	F	ASL	ASL American Sign Language	This application offers flashcards and quizzes, but it needs improvement, as too many adds appear and user interface is not appropriate	Flashcards, quizzes

Table 16 (continued)

Sr. #	Android (A)/IOS (I)/Both (B)	Paid (P)/ Free (F)/ Both (B)	Language	Application Name	Description	Best For
Other than ASL language learning apps						
15	B	F	AusLan	RIDBC Auslan Tutor Key Signs	Learn almost 150 most frequently used AusLan words and phrases. It uses photos of the hand shapes and video clips for learning the language	Learning basics of AusLan
16	A	F	IrSL	IrSL Everywhere	This app teaches Irish Sign Language. It is just a list of numbers, letters, words grouped into different categories	Irish Sign Language Learning
17	A	F	PorSL	ProDeaf Translator	Type in any phrase, and the animated man will sign it for you. The app is designed mainly for Portuguese Sign Language, but it supports ASL in beta as well. Sadly, many reviewers have stated that due to the beta status of the English side, the translations can be awkward or clunky	Portuguese, animated demonstrations
18	B	F	NZSL	NZSL Dictionary	A dictionary for New Zealand Sign Language Dictionary that contains more than 4500 videos	Learning NZSL
19	A	F	ISL	ISL Dictionary	Daily life words for the deaf community of India in the form of videos. It also has translation system that translates in 11 Indian languages	Learning and translation of ISL
20	A	F	BSL	Sign BSL	BSL sign dictionary with more than 20,000 videos. It contains letters, words, phrases of BSL	A big multimedia dictionary of BSL
21	A	F	ArgSL	LSApp	LSApp is a free application supported by Argentine Association of the Deaf to learn the basics of ArgSL. It contains validated signs and user experience	Learning basics of Argentinian SL
22	A	F	DGS	Gebärden Lernen	An online video dictionary for the German Sign Language (DGS) based on the sign language dictionary of gebaerdenlernen.de	Learning DGS
23	A	F	LSB	Hand Talk	Automatically translates the text and audio to Brazilian Sign Language	Basics of BrSL
24	A	F	InSL	Belajar Bahasa Isyarat	Sign language learning application is an application to learn sign language that starts from letters	Basics of Indonesian SL
25	A	F	TSL	I¸sareṫ Dili-Hareketli Sözlük-9500 + I¸sareṫ	Helps learn and read letters, is equipped with guessing game of picture letters, guess the picture of words and also equipped with pictures to help the deaf	Learn basics of Turkish SL
26	A	F	GrSL	Greek Sign Language	An application that offers learning Greek Sign Language with the help of video-based dictionary It has been developed by the Center of Greek Sign Language	Learning basics of Greek SL
27	A	F	MxSL	Aprende señas: Lengua de Señas Mexicana	With "Learn signs: Mexican Sign Language," you will learn more than 180 signs. Play in the 12 different categories: Alphabet, Numbers, Colors, Animals, Professions, Sports, Greetings, Places, Dates, Clothes, Family and Food	Dictionary of MxSL in different categories

Table 16 (continued)

Sr. #	Android (A)/IOS (I)/Both (B)	Paid (P)/ Free (F)/ Both (B)	Language	Application Name	Description	Best For
28	A	F	RSL		The application for learning the Russian Sign Language will allow you to learn the alphabet of fingerprints, gestures, the basics of grammar and common phrases of RJJ in a playful way. In total, about 1400 animated gestures and their combinations are available for study, divided into 28 thematic lessons	Learning RSL basics along with different lessons
29	A	F	JSL	Bunpo: Learn Japanese	A free app to learn basic to advance level JSL. It provides grammar explanation and lessons for learners of any levels, and contains more than 1700 example sentences, and also contains 8000 Japanese grammar quiz questions. Lastly, it comes with English-Japanese translations	Learning Japanese SL beginner to advance. Also offers translation
30	A	P	Multilingual	Spread Signs	Acting as a searchable video dictionary. If you spring for the in-app purchase, though, you can delve into the alphabet, favorite words, and access a section designed for children	Multiple languages, slowed-down videos of real people signing
31	B	B	ASL/ Chinese	Baby Sign Language Dictionary	It helps babies learn sign language gestures for routine words, animals, action words, feelings words. It also provides more than 40 videos. It assesses the learning with the help of quizzes like assessments	English, Simplified Chinese, traditional Chinese
Domain specific applications						
32	B	F	ASL	JW Language	It is an official app developed by JW Language which helps officials and language learners improve their SL skills with the help of while working for ministry and congregation meeting	ASL learning for ministry and congregation meetings
33	I	F	ASL	WeSign Basic	An app for parents to communicate with their deaf children by asking routine questions related to schools	App for parents to manage schooling activities
34	B	F	ArSL	Sign Language-	An application for Dubai police to learn ArSL. It also helps the Police to communicate with deaf persons in the society	Dubai Police
35	A	F	PSL	Prayer for deaf	It is an application for the deaf community to learn about different concepts and practices related to Islam	Learning Islam
36	A	F	PSL	Deaf and Mute Muslims	It is an app developed by Dawat-e-Islami, a global non-political movement for the preaching of Quran and Sunnah. This app addresses different types of special people including deaf, mute, and blind and helps them learning basics of religion Islam	Learning basics of Islam
37	B		ASL	Iglesia Ni Cristo Sign Language App	The Iglesia Ni Cristo Sign Language App is a project of the Christian Society for the Deaf under the Christian Family Organizations Office of the Iglesia Ni Cristo (Church Of Christ). It is part of the Church's continuous intensive efforts to reach out to and care for the Deaf, INC members and non-members alike	Learn about Christianity
Translation systems and interpreters						
38	A	F	ASL	Talk to deaf	If one of your friends or loved ones is deaf or just got hearing impaired it is quite hard to communicate with them. With this app you use the mobile to translate your speech to text. The deaf person can then easily read your message	Translates voice to SL gestures
39	B	F	ASL	Purple VRI	It is an on demand sign language interpretation service. It serves different scenarios where interpreters are required to communicate with the deaf persons	Translation and interpreter service

sentences. It is pertinent to mention that direct translation of gestures into a sequence of words does not result in grammatically and syntactically correct natural language sentences.

9 RQ6: taxonomy and challenges in sign language translation

9.1 Taxonomy for sign language translation research

The detailed analysis of research work conducted in the area of sign language translation reveals that it can be divided into different sub-areas. Figure 30 presents a classification that has been extracted after conducting this detailed survey in the domain of sign language translation. The taxonomy shows the following different major branches of research for sign language translation. One important aspect is that of sign language repositories and dictionaries that contain gestures in different modes including sign writing notations. The other facet involves natural to sign language translation systems which are mostly based on semantic and statistical translation models, while some machine learning- and deep learning-based translation models have also been proposed recently. A significant amount of work has been conducted on gesture recognition systems. These approaches can be categorized into the ones that involve any type of wearable hardware or specialized cameras, or the others which are purely computer vision-based approaches and solely process the videos or images without the involvement of any hardware. Furthermore, another important aspect is development of avatar technology to automatically perform the gesture, thus minimizing the need of a human interpreter. Lastly, there exist a variety of applications for the deaf community. Some of them help them to learn and practice a sign language, while others help in translation, whereas many domain-specific applications have also been developed to bridge the communication gap between the deaf and normal persons.

9.2 Open challenges in sign language translation

There are many interesting challenges associated with every distinct aspect of sign language translation discussed in this study. This section presents some of the promising future research directions for all different aspects of sign language translation shown in its taxonomy.

9.2.1 Sign language repositories

Though there exist sign language dictionaries for almost every sign language, they need to be improved in the following ways.

Standardization: Very few sign languages have standardized dictionaries. ASL is a linguistically well studied language and has standard gestures and online dictionaries. However, most of the developing countries need to work on standardization of their sign language so as to make a proper national sign language, just like the spoken languages. This will certainly help reducing the communication gap.

Sign writing notations: The dictionaries need to be augmented by providing the sign writing notations, preferably the notations that can help generating avatar-based gestures using different software systems. Recently, Wehrmeyer has proposed a new sign writing notation system [180] that incorporates international phonetic alphabets for sign languages. It claims to present a way of representing hand shapes suitable for lexicography, studying phonetic variation, and avatar programming. The proposed method has been initially tested on South African Sign Language.

Mechanism for adding words: There is a need for a body that not only helps in standardization of a sign language, but also responsible for adding gestures for new words. Particularly, the words for different technological domains need to be created and made a part of such dictionaries by a well-defined process.

9.2.2 Gesture recognition

Gesture recognition involves several open challenges in different categories.

Nature of gestures: It is one important aspect, as there are several ingredients of gestures as discussed earlier in the article. Thus, accurate recognition of single- and double-handed gestures along with their movements is certainly challenging. Apart from this, the recognition of non-manual features is not in the scope of most of the studies. Furthermore, some gestures are very similar as shown in Fig. 31, and it is very hard for gesture recognition approaches, particularly software-based approaches, to differentiate among them.

Subjects performing gestures: Every person has a different ways of performing gestures that may depend upon the proficiency of a person in the sign language; or may depend on the speed of performing gestures, size of hands,

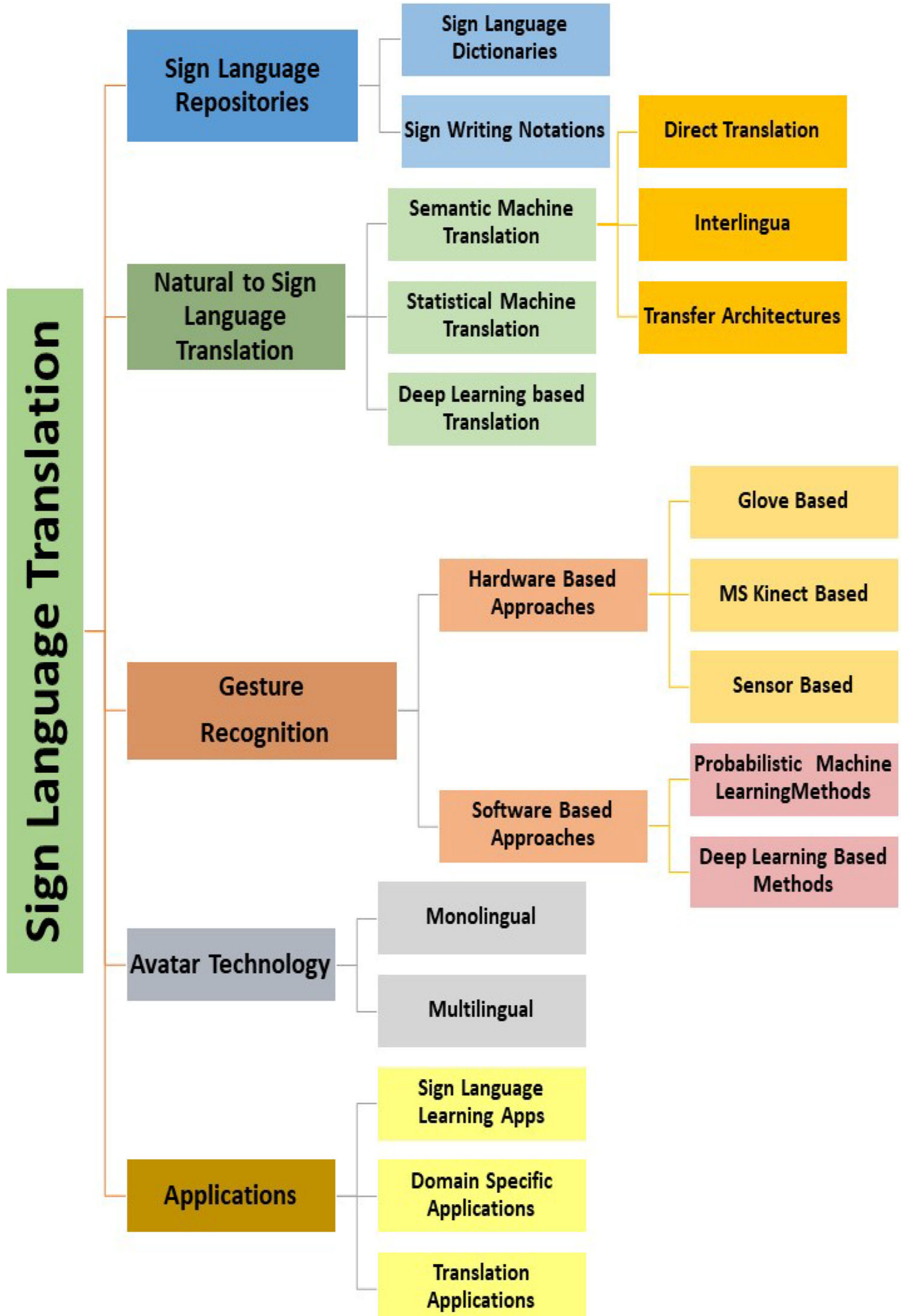


Fig. 30 Taxonomy of sign language translation system components

and may also involve other variations. This makes it challenging.

Hardware: The hardware-based approaches have many different inherent challenges. Firstly, the cost of hardware is also an important aspect. There is a need to bring the cost of hardware down so that it may become affordable for the general public to benefit from hardware-based technology to recognize gestures performed by a deaf person. Secondly, these devices need to be calibrated for different users due to the anatomical differences in different subjects, which poses questions to the accuracy of results. Similarly, number and local of sensors also play their role and offer a trade-off in terms of accuracy and cost.

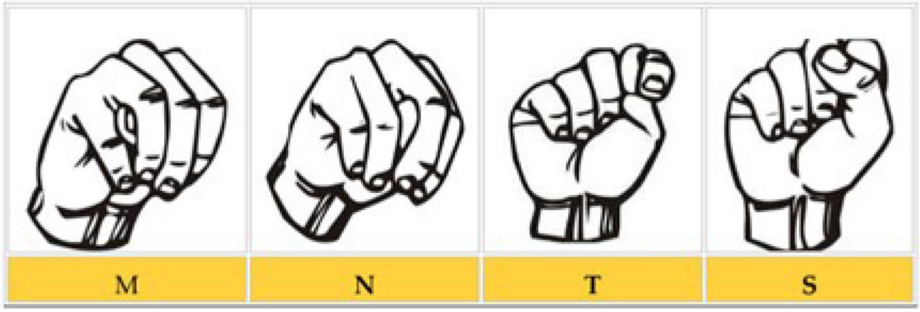


Fig. 31 Gestures for different ASL alphabets with almost similar hand shapes [17]

9.2.3 Avatar technology

The avatar technology needs to be improved in the following different ways:

More suitable avatar: The avatars need to be more like humans, and different variations should exist for children and adults. Furthermore, there is a strong need to work on the improvement in the non-manual features for the avatars so that they should be able to provide semantically meaningful non-manual cues. Realistic Transition: The generation of a gesture involves pulling complete sign executions from motion-capture libraries, but then requires to gel these executions in a sequence. One sign may end with hands in one position, while the other may require them to start from another position. This requires smoother transition of avatar.

Modeling modulations: The adjectives and adverbs are executed by modulating a noun or verb. As an example, the gesture for plane ride in ASL involves moving a certain hand shape showing the aeroplane flight, while the flight with turbulence is almost same while showing the movement bumpy. Capturing all different descriptors in the motion-capture data is not a feasible option and needs to be addressed.

Public data sets: Most of the motion-capture data sets are proprietary and are owned by different organizations and some others by different research groups. This unavailability of data sets is an impeding factor in the improvement of avatar technology.

9.2.4 General challenges and directions

Following are some general challenges and research directions that need to be addressed to improve the state of the art for sign language translation systems.

Lack of data sets: The lack of data set from all aspects is a fundamental problem and needs to be addressed using modern technologies. The sign language corpora are order of magnitude smaller than speech or other natural language corpora, as shown in Table 8. The reason is that sign languages are not typically written, and the parallel written corpora are absent as far as sign languages are concerned. Similarly, the corpora for gestures of sign language are also limited and need to be extended by involving more and more deaf subjects and experts of sign languages.

Crowd sourcing with editorial control: There is a need to engage the deaf community to develop and test translation systems. The deaf personnel can be engaged using a crowd sourcing platform to generate parallel written sign language corpora that contain natural language text and equivalent sign language text. Similarly, the deaf community can be engaged to contribute for gestures of different words using a controlled editorial process. This can help in many ways: (a) It can help gathering signs for gesture recognition with multiple signers; (b) we can gather data for regional sign languages in an easier way; (c) we can engage the deaf community while designing gestures for new words; and (d) we can engage the deaf community to help evaluating the qualitative parameters and the acceptability of the developed translation- and avatar-based system.

Multi-lingual translation applications: There is also a need to develop multilingual translation applications. This can help the deaf community from different parts of the world communicate with one another. Furthermore, it can help people translate from one sign language to the other, thus making the content available in one language useful for many others.

10 Conclusion

This article presents a systematic literature review on the state of the art of sign language translation. It presents the fundamentals of sign languages involving the types of gestures, linguistic aspects showing how a sign language differs from other natural languages. All the major ingredients required for sign language translation were discussed one after the other. A detailed discussion about the dictionaries and sign language repositories was presented. This was augmented by highlighting the importance of sign writing notations. Different variants of the two-way translation systems, i.e., natural to sign language translation and vice versa (gesture recognition), were discussed in details.

Developments in the usage of avatar technology for imitating the gestures performed by humans were also discussed. Furthermore, a variety of applications developed for learning sign language, translation applications, and some domain specific applications for different sign languages were discussed. In the end, a taxonomy of the whole sign language translation system was presented along with pertinent open challenges and future research directions in different sub-areas of sign language translation.

Acknowledgements This study is part of Ms Uzma Farooq's doctoral thesis work at Faculty of Computing, Universiti Teknologi Malaysia, Malaysia. Prof. Hussain would like to acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) - Grants Ref. EP/M026981/1, EP/T021063/1, EP/T024917/1. The authors also wish to thank the anonymous reviewers for their insightful comments and suggestions which helped improve the quality of our paper. Part of this work has been presented in a conference paper available at https://ieeexplore.ieee.org/document/8966714.

Funding There is no funding available for this research.

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Conflicts of interest The authors do not have any conflict of interest or any competing interests.

References

1. "Cybertouch—Cyberglove systems LLC" (2020) http://www.cyberglovesystems.com/cybertouch. Accessed 21 April 2020
2. "Digital assistance for sign language users" (2020) https://www.microsoft.com/en-us/research/blog/digital-assistance-for-sign-language-users/. Accessed 21 April 2020
3. "ESIGN project" (2020) https://www.sign-lang.uni-hamburg.de/esign/. Accessed 21 April 2020
4. "National association of the deaf—NAD" (2020) https://www.nad.org/. Accessed 21 April 2020
5. "Accelerometer sensor—an overview—ScienceDirect topics" (2020) https://www.sciencedirect.com/topics/engineering/accelerometer-sensor. Accessed 21 April 2020
6. "American Sign Language video dataset" (2020) http://vlm1.uta.edu/~athitsos/asl_lexicon/. Accessed 21 April 2020
7. "American Sign Language video dataset" (2020) http://csr.bu.edu/asl/asllvd/annotate/index-cvpr4hb08-dataset.html. Accessed 21 April 2020
8. "Deafness and hearing loss" (2020) https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. Accessed 21 April 2020
9. "Directorate general of special education, Islamabad, Pakistan" (2020) http://dgse.gov.pk/html/index.html. Accessed 21 April 2020
10. "Electromyography (EMG), Johns Hopkins Medicine" (2020) https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electromyography-emg. Accessed 21 April 2020
11. "Image data set—automatic sign language detection" (2020) https://sites.google.com/site/autosignlan/source/image-data-set. Accessed 21 April 2020
12. "Leap motion" (2020) https://www.leapmotion.com/. Accessed 21 April 2020
13. "National Association of the Deaf, India" (2020) http://nadindia.org/. Accessed 21 April 2020
14. "Sign language digits dataset" (2020) https://kaggle.com/ardamavi/sign-language-digits-dataset. Accessed 21 April 2020
15. "ASL: fingerspelling (lifeprint.com)" (2020) https://itunes.apple.com/us/app/asl-fingerspelling-lifeprint-com/id605558017?mt=8. Accessed 21 April 2020
16. Agarwal A, Thakur MK (2013) Sign language recognition using Microsoft kinect. In: 2013 sixth international conference on contemporary computing (IC3). IEEE, pp 181–185
17. Ahmed MA, Zaidan BB, Zaidan AA, Salih MM, Lakulu MM (2018) A review on systems based sensory gloves for sign language recognition state of the art between 2007 and 2017. Sensors 18(7):2208
18. Al-khazraji S, Berke L, Kafle S, Yeung P, Huenerfauth M (2018) Modeling the speed and timing of American Sign Language to generate realistic animations. In: Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility, pp 259–270
19. Almeida SG, Guimarães FG, Ramírez JA (2014) Feature extraction in Brazilian Sign Language recognition based on phonological structure and using RGB-D sensors. Expert Syst Appl 41(16):7259–7271
20. AlShammari A, Alsumait A, Faisal M (2018) Building an interactive e-learning tool for deaf children: interaction design process framework. In: 2018 IEEE conference on e-learning, e-management and e-services (IC3e). IEEE, pp 85–90
21. Ambar R, Fai CK, Abd Wahab MH, Jamil MM, Ma'radzi AA (2018) Development of a wearable device for sign language recognition. In: Journal of physics: conference series, vol 1019. IOP Publishing, p 012017
22. Andrei S, Osborne L, Smith Z (2013) Designing an American Sign Language avatar for learning computer science concepts for deaf or hard-of-hearing students and deaf interpreters. J Educ Multimedia Hypermedia 22(3):229–242
23. Antony PJ (2013) Machine translation approaches and survey for Indian Languages. Int J Comput Linguist Chin Lang Process 18(1):47–78
24. Athira PK, Sruthi CJ, Lijiya A (2019) A signer independent sign language recognition with co-articulation elimination from live videos: an Indian scenario. J King Saud Univ-Comput Inf Sci. https://doi.org/10.1016/j.jksuci.2019.05.002
25. Avola D, Bernardi M, Cinque L, Foresti GL, Massaroni C (2018) Exploiting recurrent neural networks and leap motion controller for the recognition of sign language and semaphoric hand gestures. IEEE Trans Multimedia 21(1):234–245
26. Bangham JA, Cox SJ, Elliott R, Glauert JR, Marshall I, Rankov S, Wells M (2000) Virtual signing: capture, animation, storage and transmission-an overview of the visicast project
27. Bastos IL, Angelo MF, Loula AC (2015) Recognition of static gestures applied to Brazilian Sign Language (libras). In: 2015 28th SIBGRAPI conference on graphics, patterns and images. IEEE, pp 305–312
28. Bonham ME (2015) English to ASL gloss machine translation
29. Bouzid Y, Jemni M (2013) An animated avatar to interpret signwriting transcription. In: 2013 international conference on electrical engineering and software applications. IEEE, pp 1–5
30. Bouzid Y, Jemni M et al (2016) The effect of avatar technology on sign writing vocabularies acquisition for deaf learners. In: 2016 IEEE 16th international conference on advanced learning technologies (ICALT). IEEE, pp 441–445

31. Bragg D, Koller O, Bellard M, Berke L, Boudreault P, Braffort A, Caselli N, Huenerfauth M, Kacorri H, Verhoef T et al (2019) Sign language recognition, generation, and translation: an interdisciplinary perspective. In: The 21st international ACM SIGACCESS conference on computers and accessibility, pp 16–31
32. Brock H, Farag I, Nakadai K (2020) Recognition of non-manual content in continuous Japanese Sign Language. Sensors 20(19):5621
33. Cate H, Hussain Z (2017) Bidirectional American Sign Language to English translation. arXiv preprint, arXiv:1701.02795
34. Ceni A, Ashwin P, Livi L (2019) Interpreting recurrent neural networks behaviour via excitable network attractors. Cogn Comput 12:1–27
35. Chai X, Li G, Lin Y, Xu Z, Tang Y, Chen X, Zhou M (2013) Sign language recognition and translation with kinect. In: IEEE conference on AFGR, vol 655, p 4
36. Chai X, Wang H, Chen X (2014) The devisign large vocabulary of Chinese Sign Language database and baseline evaluations. Technical report VIPL-TR-14-SLR-001. Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS
37. Chiriac IA, Stoicu-Tivadar L, Podoleanu E (2015) Comparing video and avatar technology for a health education application for deaf people. In: MIE, pp 516–520
38. Chiu YH, Wu CH, Su HY, Cheng CJ (2006) Joint optimization of word alignment and epenthesis generation for Chinese to Taiwanese sign synthesis. IEEE Trans Pattern Anal Mach Intell 29(1):28–39
39. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint, arXiv:1406.1078.
40. Chuan CH, Regina E, Guardino C (2014) American Sign Language recognition using leap motion sensor. In: 2014 13th international conference on machine learning and applications. IEEE, pp 541–544
41. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint, arXiv:1412.3555
42. Camgoz NC, Hadfield S, Koller O, Ney H, Bowden R (2018) Neural sign language translation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7784–7793
43. Cormier K, Fenlon J, Johnston T, Rentelis R, Schembri A, Rowley K, Adam R, Woll B (2012) From corpus to lexical database to online dictionary: issues in annotation of the BSL corpus and the development of BSL SignBank. In: Proceedings of the 5th workshop on the representation and processing of sign languages: interactions between corpus and lexicon, international conference on language resources and evaluation, LREC, pp 7–12
44. Cox S, Lincoln M, Tryggvason J, Nakisa M, Wells M, Tutt M, Abbott S (2002) Tessa, a system to aid communication with deaf people. In: Proceedings of the fifth international ACM conference on assistive technologies, pp 205–212
45. Cui R, Liu H, Zhang C (2019) A deep neural framework for continuous sign language recognition by iterative training. IEEE Trans Multimedia 21(7):1880–1891
46. d'Armond LS (2002) Representation of American Sign Language for machine translation. Ph.D. thesis, Georgetown University
47. Das A, Yadav L, Singhal M, Sachan R, Goyal H, Taparia K, Gulati R, Singh A, Trivedi G (2016) Smart glove for sign language communications. In: 2016 international conference on accessibility to digital world (ICADW). IEEE, pp 27–31
48. Debevc M, Milošević D, Kožuh I (2015) A comparison of comprehension processes in sign language interpreter videos with or without captions. PLoS ONE 10(5):e0127577
49. Di Mascio T, Gennari R (2008) An intelligent visual dictionary for Italian Sign Language. J Web Eng 7(4):318–338
50. Dipietro L, Sabatini AM, Dario P (2008) A survey of glove-based systems and their applications. IEEE Trans Syst Man Cybern Part C (Appl Rev) 38(4):461–482
51. Ebling S, Glauert J (2016) Building a Swiss German Sign Language avatar with JASigning and evaluating it among the deaf community. Univ Access Inf Soc 15(4):577–587
52. El-Bendary N, Zawbaa HM, Daoud MS, Hassanien AE, Nakamatsu K (2010) Arslat: Arabic sign language alphabets translator. In: 2010 international conference on computer information systems and industrial management applications (CISIM). IEEE, pp 590–595
53. Elghoul MJ (2007) An avatar based approach for automatic interpretation of text to sign language. Chall Assist Technol: AAATE 07(20):266
54. Elliott R, Glauert JR, Kennaway JR, Marshall I (2000) The development of language processing support for the ViSiCAST project. In: Proceedings of the fourth international ACM conference on assistive technologies, pp 101–108
55. Elliott R, Glauert JR, Kennaway JR, Marshall I, Safar E (2008) Linguistic modelling and language-processing technologies for avatar-based sign language presentation. Univ Access Inf Soc 6(4):375–391
56. Elons AS, Ahmed M, Shedid H, Tolba MF (2014) Arabic Sign Language recognition using leap motion sensor. In: 2014 9th international conference on computer engineering & systems (ICCES). IEEE, pp 368–373
57. Eźlakowski W (2020) Grammar of Polish Sign Language as compared to grammar of polish language: selected themes. Sign Lang Stud 20(3):518–532
58. Fadlilah U, Wismoyohadi D, Mahamad AK, Handaga B (2019) Bisindo information system as potential daily sign language learning. In: AIP conference proceedings, vol 2114. AIP Publishing LLC, p 060021
59. Farooq U, Asmat A, Rahim MS, Khan NS, Abid A (2019) A comparison of hardware based approaches for sign language gesture recognition systems. In: 2019 international conference on innovative computing (ICIC). IEEE, pp 1–6
60. Filhol M (2009) Zebedee: a lexical description model for sign language synthesis. Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI). LIMSICNRS, Orsay
61. Firdaus M, Golchha H, Ekbal A, Bhattacharyya P (2020) A deep multi-task model for dialogue act classification, intent detection and slot filling. Cogn Comput, 1–20
62. Forster J, Schmidt C, Koller O, Bellgardt M, Ney H (2014) Extensions of the sign language recognition and translation corpus RWTH-PHOENIX-weather. In: LREC, pp 1911–1916
63. Fuertes JL, González ÁL, Mariscal G, Ruiz C (2006) Bilingual sign language dictionary. In: International conference on computers for handicapped persons. Springer, pp 599–606
64. Gao F, Huang T, Sun J, Wang J, Hussain A, Yang E (2019) A new algorithm for SAR image target recognition based on an improved deep convolutional neural network. Cogn Comput 11(6):809–824
65. Garje GV, Kharate GK (2013) Survey of machine translation systems in India. Int J Nat Lang Comput (IJNLC) 2:47–67
66. Geetha M, Manjusha C, Unnikrishnan P, Harikrishnan R (2013) A vision based dynamic gesture recognition of Indian Sign Language on kinect based depth images. In: 2013 international conference on emerging trends in communication, control,

signal processing and computing applications (C2SPCA), pp 1–7. IEEE
67. Gibet S, Courty N, Duarte K, Naour TL (2011) The SignCom system for data-driven animation of interactive virtual signers: methodology and evaluation. ACM Trans Interact Intell Syst (TiiS) 1(1):1–23
68. Goyal L, Goyal V (2016) Text to sign language translation system: a review of literature. Int J Synth Emot (IJSE) 7(2):62–77
69. Grieve-Smith AB (2001) Signsynth: a sign language synthesis application using web3d and perl. In: International gesture workshop. Springer, pp 134–145
70. Grimes GJ (1983) Digital data entry glove interface device, 8 Nov 1983. US Patent 4,414,537
71. Halawani SM, Zaitun AB (2012) An avatar based translation system from Arabic speech to Arabic Sign Language for deaf people. Int J Inf Sci Educ 2(1):13–20
72. Halim Z, Abbas G (2015) A kinect-based sign language hand gesture recognition system for hearing-and speech-impaired: a pilot study of Pakistani Sign Language. Assist Technol 27(1):34–43
73. Hanke T, König L, Wagner S, Matthes S (2010) DGS corpus & Dicta-sign: the Hamburg studio setup. 4th workshop on the representation and processing of sign languages: corpora and sign language technologies (CSLT 2010), Valletta, Malta, pp 106–110
74. Haydar J, Dalal B, Hussainy S, El Khansa L, Fahs W (2012) ASL fingerspelling translator glove. Int J Comput Sci Issues (IJCSI) 9(6):254
75. Heloir A, Kipp M (2010) Real-time animation of interactive agents: specification and realization. Appl Artif Intell 24(6):510–529
76. Huang J, Zhou W, Li H, Li W (2018) Attention-based 3D-CNNS for large-vocabulary sign language recognition. IEEE Trans Circuits Syst Video Technol 29(9):2822–2832
77. Huang J, Zhou W, Zhang Q, Li H, Li W (2018) Video-based sign language recognition without temporal segmentation. In: Thirty second AAAI conference on artificial intelligence
78. Huenerfauth M, Kacorri H (2014) Release of experimental stimuli and questions for evaluating facial expressions in animations of American Sign Language. In: Proceedings of the 6th workshop on the representation and processing of sign languages: beyond the manual channel, the 9th international conference on language resources and evaluation (LREC 2014), Reykjavik, Iceland
79. Huenerfauth M, Marcus M, Palmer M (2006) Generating American Sign Language classifier predicates for English-to-ASL machine translation. Ph.D. thesis, University of Pennsylvania.
80. Huenerfauth MP (2003) American Sign Language natural language generation and machine translation systems. Technical report. Computer and Information Sciences, University of Pennsylvania
81. Hutchinson J (2012) Literature review: analysis of sign language notations for parsing in machine translation of SASL. Ph.D. thesis, Rhodes University, South Africa
82. Ibrahim NB, Selim MM, Zayed HH (2018) An automatic Arabic Sign Language Recognition System (ArSLRS). J King Saud Univ Comput Inf Sci 30(4):470–477
83. Jalal MA, Chen R, Moore RK, Mihaylova L (2018) American Sign Language posture understanding with deep neural networks. In: 2018 21st international conference on information fusion (FUSION). IEEE, pp 573–579
84. Jin CM, Omar Z, Jaward MH (2016) A mobile application of American Sign Language translation via image processing algorithms. In: 2016 IEEE Region 10 Symposium (TENSYMP). IEEE, pp 104–109
85. Jingqiu W, Ting Z (2014) An arm-based embedded gesture recognition system using a data glove. In: The 26th Chinese control and decision conference (2014 CCDC). IEEE, pp 1580–1584
86. Joze HR, Koller O (2018) Ms-asl: a large-scale data set and benchmark for understanding American Sign Language. arXiv preprint, arXiv:1812.01053
87. Kacorri H, Huenerfauth M (2016) Continuous profile models in ASL syntactic facial expression synthesis. In: Proceedings of the 54th annual meeting of the association for computational linguistics (volume 1: long papers), pp 2084–2093
88. Kacorri H, Huenerfauth M, Ebling S, Patel K, Menzies K, Willard M (2017) Regression analysis of demographic and technology-experience factors influencing acceptance of sign language animation. ACM Trans Access Comput (TACCESS) 10(1):1–33
89. Kanwal K, Abdullah S, Ahmed YB, Saher Y, Jafri AR (2014) Assistive glove for Pakistani Sign Language translation. In: 17th IEEE international multi topic conference 2014. IEEE, pp 173–176
90. Karbasi M, Zabidi A, Yassin IM, Waqas A, Bhatti Z (2017) Malaysian Sign Language dataset for automatic sign language recognition system. J Fundam Appl Sci 9(4S):459–474
91. Kennaway JR, Glauert JR, Zwitserlood I (2007) Providing signed content on the internet by synthesized animation. ACM Trans Comput-Hum Interact (TOCHI) 14(3):15-es
92. Khan NS, Abid A, Abid K (2020) A novel natural language processing (NLP)–based machine translation model for English to Pakistan Sign Language translation. Cogn Comput 12:748–765
93. Kim J, Wagner J, Rehm M, André E (2008) Bichannel sensor fusion for automatic sign language recognition. In: 2008 8th IEEE international conference on automatic face & gesture recognition. IEEE, pp 1–6
94. Kim T, Keane J, Wang W, Tang H, Riggle J, Shakhnarovich G, Brentari D, Livescu K (2017) Lexicon-free fingerspelling recognition from video: Data, models, and signer adaptation. Comput Speech Lang 46:209–232
95. Kipp M, Heloir A, Nguyen Q (2011) Sign language avatars: animation and comprehensibility. In: International workshop on intelligent virtual agents. Springer, pp 113–126
96. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering
97. Koller O, Camgoz NC, Ney H, Bowden R (2019) Weakly supervised learning with multi-stream CNN-LSTM-HMMS to discover sequential parallelism in sign language videos. IEEE Trans Pattern Anal Mach Intell 42(9):2306–2320
98. Koller O, Zargaran S, Ney H (2017) Re-sign: re-aligned end-to-end sequence modelling with deep recurrent CNN-HMMS. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4297–4305
99. Kouremenos D, Ntalianis K, Kollias S (2018) A novel rule based machine translation scheme from Greek to Greek Sign Language: production of different types of large corpora and language models evaluation. Comput Speech Lang 51:110–135
100. Kristoffersen JH, Troelsgård T (2012) The electronic lexicographical treatment of sign languages: the Danish Sign Language Dictionary. Oxford University Press, Oxford
101. Krňoul Z, Kanis J, Železný M, Müller L (2007) Czech text-to-sign speech synthesizer. In: International workshop on machine learning for multimodal interaction. Springer, pp 180–191
102. Kumar P, Gauba H, Roy PP, Dogra DP (2017) Coupled HMM-based multi-sensor data fusion for sign language recognition. Pattern Recogn Lett 86:1–8

103. Lang S, Block M, Rojas R (2012) Sign language recognition using kinect. In: International conference on artificial intelligence and soft computing. Springer, pp 394–402
104. Lee BG, Lee SM (2017) Smart wearable hand device for sign language interpretation system with sensors fusion. IEEE Sens J 18(3):1224–1232
105. Li Y, Chen X, Zhang X, Wang K, Wang ZJ (2012) A sign-component-based framework for Chinese Sign Language recognition using accelerometer and SEMG data. IEEE Trans Biomed Eng 59(10):2695–2704
106. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 62(10):e1–e34
107. Liu Y, Vong CM, Wong PK (2017) Extreme learning machine for huge hypotheses re-ranking in statistical machine translation. Cogn Comput 9(2):285–294
108. Liwicki S, Everingham M (2009) Automatic recognition of finger spelled words in British Sign Language. In: 2009 IEEE computer society conference on computer vision and pattern recognition workshops. IEEE, pp 50–57
109. Lokhande P, Prajapati R, Pansare S (2015) Data gloves for sign language recognition system. Int J Comput Appl 975:8887
110. Lu D, Yu Y, Liu H (2016) Gesture recognition using data glove: An extreme learning machine method. In: 2016 IEEE international conference on robotics and biomimetics (ROBIO). IEEE, pp 1349–1354
111. Lu P, Huenerfauth M (200) Collecting a motion-capture corpus of American Sign Language for data-driven generation research. In: Proceedings of the NAACL HLT 2010 workshop on speech and language processing for assistive technologies. Association for Computational Linguistics, pp 89–97
112. Lun R, Zhao W (2015) A survey of applications and human motion recognition with microsoft kinect. Int J Pattern Recognit Artif Intell 29(05):1555008
113. Luqman H, Mahmoud SA (2019) Automatic translation of Arabic Text-to-Arabic Sign Language. Univ Access Inf Soc 18(4):939–951
114. Luqman H, Mahmoud SA et al (2017) Transform-based Arabic Sign Language recognition. Procedia Comput Sci 117:2–9
115. Ma J, Gao W, Wu J, Wang C (2000) A continuous Chinese Sign Language recognition system. In: Proceedings fourth IEEE international conference on automatic face and gesture recognition (Cat. No. PR00580). IEEE, pp 428–433
116. Marschark M, Sapere P, Convertino C, Seewagen R, Maltzen H (2004) Comprehension of sign language interpreting: deciphering a complex task situation. Sign Lang Stud 4(4):345–368
117. Martin PM, Belhe S, Mudliar S, Kulkarni M, Sahasrabudhe S (2013) An Indian Sign Language (ISL) corpus of the domain disaster message using avatar. In: Proceedings of the third international symposium in sign language translations and technology (SLTAT-2013), pp 1–4
118. Martínez AM, Wilbur RB, Shay R, Kak AC (2002) Purdue RVL-SLLL ASL database for automatic recognition of American Sign Language. In: Proceedings. Fourth IEEE international conference on multimodal interfaces. IEEE, pp 167–172
119. Mehdi SA, Khan YN (2002) Sign language recognition using sensor gloves. In: Proceedings of the 9th international conference on neural information processing, 2002. ICONIP'02, vol 5. IEEE, pp 2204–2206
120. Mittal A, Kumar P, Roy PP, Balasubramanian R, Chaudhuri BB (2019) A modified LSTM model for continuous sign language recognition using leap motion. IEEE Sens J 19(16):7056–7063
121. Mocialov B, Hastie H, Turner G (2018) Transfer learning for British Sign Language modelling. In: Proceedings of the fifth workshop on NLP for similar languages, varieties and dialects (VarDial 2018), pp 101–110
122. Mohandes MA (2013) Recognition of two-handed Arabic signs using the cyberglove. Arab J Sci Eng 38(3):669–677
123. Nguyen TBD, Phung TN, Vu TT (2018) A rule-based method for text shortening in Vietnamese Sign Language translation. In: Bhateja V, Nguyen B, Nguyen N, Satapathy S, Le DN (eds) Information systems design and intelligent applications. Springer, Singapore, pp 655–662
124. Niewiadomski R, Bevacqua E, Mancini M, Pelachaud C (2009) Greta: an interactive expressive ECA system. In: Proceedings of the 8th international conference on autonomous agents and multiagent systems, vol 2, pp 1399–1400
125. Nisar S, Tariq M, Adeel A, Gogate M, Hussain A (2019) Cognitively inspired feature extraction and speech recognition for automated hearing loss testing. Cogn Comput 11(4):489–502
126. Othman A, Jemni M (2011) Statistical sign language machine translation: from English written text to American Sign Language gloss. arXiv preprint, arXiv:1112.0168
127. Othman A, Jemni M (2019) Designing high accuracy statistical machine translation for sign language using parallel corpus: case study English and American Sign Language. J Inf Technol Res (JITR) 12(2):134–158
128. Otter DW, Medina JR, Kalita JK (2018) A survey of the usages of deep learning in natural language processing. arXiv preprint, arXiv:1807.10854.
129. Oz C, Leu MC (2005) Recognition of finger spelling of American Sign Language with artificial neural network using position/orientation sensors and data glove. In: International symposium on neural networks. Springer, pp 157–164
130. Ozcan T, Basturk A (2019) Transfer learning-based convolutional neural networks with heuristic optimization for hand gesture recognition. Neural Comput Appl 31(12):8955–8970
131. Parvini F, McLeod D, Shahabi C, Navai B, Zali B, Ghandeharizadeh S (2009) An approach to glove-based gesture recognition. In: International conference on human-computer interaction. Springer, pp 236–245
132. Pizzolato EB, dos Santos Anjo M, Pedroso GC (2010) Automatic recognition of finger spelling for libras based on a two-layer architecture. In: Proceedings of the 2010 ACM symposium on applied computing, pp 969–973
133. Platts JT (1884) A dictionary of Urdu, classical Hindi, and English. H. Milford, Oxford
134. Poria S, Soon OY, Liu B, Bing L (2020) Affect recognition for multimodal natural language processing. Cogn Comput 13:1–2
135. Porta J, López-Colino F, Tejedor J, Colás J (2014) A rule-based translation from written Spanish to Spanish Sign Language glosses. Comput Speech Lang 28(3):788–811
136. Potter LE, Araullo J, Carter L (2013) The leap motion controller: a view on sign language. In: Proceedings of the 25th Australian computer–human interaction conference: augmentation, application, innovation, collaboration, pp 175–178
137. Praveen N, Karanth N, Megha MS (2014) Sign language interpreter using a smart glove. In: 2014 international conference on advances in electronics computers and communications. IEEE, pp 1–5
138. Prillwitz S (1990) The long road towards bilingualism of the deaf in the German-speaking area. Sign Lang Res Appl: Int Stud Sign Lang Commun Deaf 13:13–27
139. Pu J, Zhou W, Li H (2019) Iterative alignment network for continuous sign language recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4165–4174

140. Pugeault N, Bowden R (2011) Spelling it out: real-time ASL fingerspelling recognition. In: 2011 IEEE international conference on computer vision workshops (ICCV workshops). IEEE, pp 1114–1119
141. Rabby G, Azad S, Mahmud M, Zamli KZ, Rahman MM (2020) TeKET: a tree-based unsupervised keyphrase extraction technique. Cogn Comput 12:1–23
142. Rabiner L, Juang B (1986) An introduction to hidden Markov models. IEEE ASSP Mag 3(1):4–16
143. Rad AE, Rahim MS, Kolivand H, Amin IB (2017) Morphological region-based initial contour algorithm for level set methods in image segmentation. Multimedia Tools Appl 76(2):2185–2201
144. Rakhmadi A, Othman NZ, Bade A, Rahim MS, Amin IM (2010) Connected component labeling using components neighbors scan labeling approach. J Comput Sci 6(10):1099
145. Rao GA, Kishore PV (2018) Selfie video based continuous Indian Sign Language recognition system. Ain Shams Eng J 9(4):1929–1939
146. Razmjooy N, Estrela VV, Loschi HJ (2019) A study on metaheuristic based neural networks for image segmentation purposes. In: Memon QA, Khoja SA (eds) Data science theory, analysis and applications. Taylor and Francis, Abingdon. https://doi.org/10.1201/9780429263798-2
147. Russell M, Kavanaugh M, Master J, Higgins J, Hoffmann T (2009) Computer-based signing accommodations: Comparing a recorded human with an avatar. J Appl Test Technol 10(3):1–21
148. Saldaña González G, Cerezo Sánchez J, Bustillo Díaz MM, Ata PA (2018) Recognition and classification of sign language for Spanish. Computación y Sistemas 22(1):271–277
149. San-Segundo R, Montero JM, Córdoba R, Sama V, Fernández F, D'haro LF, López-Ludena V, Sánchez D, García A (2012) Design, development and field evaluation of a Spanish into sign language translation system. Pattern Anal Appl 15(2):203–224
150. Scassellati B, Brawer J, Tsui K, Nasihati Gilani S, Malzkuhn M, Manini B, Stone A, Kartheiser G, Merla A, Shapiro A et al (2018) Teaching language to deaf infants with a robot and a virtual human. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–13
151. Segouat J, Braffort A (2009) Toward the study of sign language coarticulation: methodology proposal. In: 2009 second international conferences on advances in computer–human interactions. IEEE, pp 369–374
152. Shohieb SM (2019) A gamified e-learning framework for teaching mathematics to Arab deaf students: (supporting an acting Arabic Sign Language avatar). Ubiquitous Learn: Int J 12(1):55–70
153. Shukor AZ, Miskon MF, Jamaluddin MH, bin Ali F, Asyraf MF, bin Bahar MB et al (2015) A new data glove approach for Malaysian Sign Language detection. Procedia Comput Sci 76:60–67
154. Singha J, Das K (2015) Automatic Indian Sign Language recognition for continuous video sequence. ADBU J Eng Technol 2(1):0021105
155. Smith R, Nolan B (2013) Manual evaluation of synthesised sign language avatars. In: Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility, pp 1–2
156. Solanki Krunal M (2013) Indian Sign Languages using flex sensor glove. Int J Eng Trends Technol 4(6):2478–2480
157. Sreelekha S (2017) Statistical vs rule based machine translation; a case study on Indian language perspective. arXiv preprint, arXiv:1708.04559
158. Starner T, Pentland A (1997) Real-time American Sign Language recognition from video using hidden Markov models. In: Motion-based recognition. Springer, pp 227–243
159. Starner T, Weaver J, Pentland A (1998) Real-time American Sign Language recognition using desk and wearable computer based video. IEEE Trans Pattern Anal Mach Intell 20(12):1371–1375
160. Stokoe WC Jr (2005) Sign language structure: an outline of the visual communication systems of the American deaf. J Deaf Stud Deaf Educ 10(1):3–37
161. Stoll S, Camgoz NC, Hadfield S, Bowden R (2020) Text2sign: towards sign language production using neural machine translation and generative adversarial networks. Int J Comput Vis 128(4):891–908
162. Sun C, Zhang T, Bao BK, Xu C, Mei T (2013) Discriminative exemplar coding for sign language recognition with kinect. IEEE Trans Cybern 43(5):1418–1428
163. Sutskever I, Vinyals O, Le QV (2014) Sequence to sequence learning with neural networks. In: Advances in neural information processing systems, pp 3104–3112
164. Suzuki E, Suzuki T, Kakihana K (2006) On the web trilingual sign language dictionary to learn the foreign sign language without learning a target spoken language. In: LREC, pp 2307–2310
165. Swathi S, Jayashree LS (2019) Machine translation using deep learning: A comparison. In: International conference on artificial intelligence, smart grid and smart city applications, pp 389–395. Springer
166. Swee TT, Ariff AK, Salleh SH, Seng SK, Huat LS (2007) Wireless data gloves Malay Sign Language recognition system. In: 2007 6th international conference on information, communications & signal processing. IEEE, pp 1–4
167. Taljard E, Prinsloo D (2019) African language dictionaries for children—a neglected genre. Lexikos 29(1):199–223
168. Tao W, Lai ZH, Leu MC, Yin Z (2018) American Sign Language alphabet recognition using leap motion controller. In: Proceedings of the 2018 institute of industrial and systems engineers annual conference (IISE 2018)
169. Targon V (2016) Learning the semantics of notational systems with a semiotic cognitive automaton. Cogn Comput 8(4):555–576
170. Tripathi K, Nandi NB (2015) Continuous Indian Sign Language gesture recognition and sentence formation. Procedia Computer Science 54:523–531
171. Tumsri J, Kimpan W (2017) Thai Sign Language translation using leap motion controller. In: Proceedings of the international multiconference of engineers and computer scientists, pp 46–51
172. Veale T, Conway A, Collins B (1998) The challenges of cross-modal translation: English-to-sign-language translation in the Zardoz system. Mach Transl 13(1):81–106
173. Verma VK, Srivastava S (2018) Toward machine translation linguistic issues of Indian Sign Language. In: Speech and language processing for human–machine communications. Springer, pp 129–135
174. Villa-Angulo R, Hidalgo-Silva H (2005) A wearable neural interface for real time translation of Spanish deaf sign language to voice and writing. J Appl Res Technol 3(3):169–186
175. Vinayagamoorthy V, Glancy M, Ziegler C, Schäffer R (2019) Personalising the TV experience using augmented reality: An exploratory study on delivering synchronised sign language interpretation. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–12
176. Von Agris U, Kraiss KF (2007) Towards a video corpus for signer-independent continuous sign language recognition. In: Gesture in human–computer interaction and simulation, Lisbon, Portugal, May 2007
177. Wadhawan A, Kumar P (2020) Deep learning-based sign language recognition system for static signs. Neural Comput Appl 32:1–12

178. Wang P, Song Q, Han H, Cheng J (2016) Sequentially supervised long short-term memory for gesture recognition. Cogn Comput 8(5):982–991
179. Wehrmeyer E (2019) A corpus for signed language interpreting research. Interpreting 21(1):62–90
180. Wehrmeyer E (2019) Rethinking handshape: a new notation system for sign language. Sign Lang Linguist 22(1):83–111
181. Wolfe R, McDonald J, Davidson MJ, Frank C (2007) Using an animation-based technology to support reading curricula for deaf elementary schoolchildren. In: The 22nd annual international technology & persons with disabilities conference, pp 1–8
182. Woll B, Sutton-Spence R, Elton F (2001) Multilingualism: the global approach to sign languages. Sociolinguist Sign Lang 8:32
183. Wu J, Tian Z, Sun L, Estevez L, Jafari R (2015) Real-time American Sign Language recognition using wrist-worn motion and surface EMG sensors. In: 2015 IEEE 12th international conference on wearable and implantable body sensor networks (BSN). IEEE, pp 1–6
184. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, Macherey W, Krikun M, Cao Y, Gao Q, Macherey K, et al (2016) Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint, arXiv:1609.08144
185. Xu C, Wang G (2019) Bidirectional cognitive computing model for uncertain concepts. Cogn Comput 11(5):613–629
186. Yang H-D (2015) Sign language recognition with the kinect sensor based on conditional random fields. Sensors 15(1):135–147
187. Young T, Hazarika D, Poria S, Cambria E (2018) Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag 13(3):55–75
188. Young T, Hazarika D, Poria S, Cambria E (2011) American Sign Language recognition with the kinect. In: Proceedings of the 13th international conference on multimodal interfaces, pp 279–286
189. Zahedi M, Dreuw P, Rybach D, Deselaers T, Ney H (2006) Continuous sign language recognition-approaches from speech recognition and available data resources. In: Second workshop on the representation and processing of sign languages: lexicographic matters and didactic scenarios, pp 21–24
190. Zhang T, Yang Y, Zeng Y, Zhao Y (2020) Cognitive template clustering improved LineMod for efficient multi-object pose estimation. Cogn Comput 12:834–843
191. Zhang X, Chen X, Li Y, Lantz V, Wang K, Yang J (2011) A framework for hand gesture recognition based on accelerometer and EMG sensors. IEEE Trans Syst Man Cybernet-Part A: Syst Hum 41(6):1064–1076
192. Zhao L, Kipper K, Schuler W, Vogler C, Badler N, Palmer M (2000) A machine translation system from English to American Sign Language. In: Conference of the Association for Machine Translation in the Americas. Springer, pp 54–67
193. Zhong G, Jiao W, Gao W, Huang K (2020) Automatic design of deep networks with neural blocks. Cogn Comput 12(1):1–12
194. Zhu G, Zhang L, Shen P, Song J (2017) Multimodal gesture recognition using 3-D convolution and convolutional LSTM. IEEE Access 5:4517–4524