Neural Computing and Applications (2025) 37:841–857
https://doi.org/10.1007/s00521-024-10395-9

REVIEW

Sign language interpretation using machine learning and artificial intelligence

Fatma M. Najib[1]

Received: 22 August 2023 / Accepted: 20 August 2024 / Published online: 18 November 2024

Abstract
Sign language is the only way for deaf and mute people to represent their needs and feelings. Most of non-deaf-mute people do not understand sign language, which leads to many difficulties for deaf-mutes' communication in their social life. Sign language interpretation systems and applications get a lot of attention in the recent years. In this paper, we review sign language recognition and interpretation studies based on machine learning, image processing, artificial intelligence, and animation tools. The two reverse processes for sign language interpretation are illustrated. This study discusses the recent research on sign language translation to text and speech with the help of hand gestures, facial expressions interpretation, and lip reading. Also, state of the art in speech to sign language translation is discussed. In addition, some of the popular and highly rated Android and Apple mobile applications that facilitate disabled people communication are presented. This paper clarifies and highlights the recent research and real used applications for deaf-mute people help. This paper tries to provide a link between research proposals and real applications. This link can help covering any gap or non-handled functionalities in the real used applications. Based on our study, we introduce a proposal involves set of functionalities/options that separately introduced and discussed by the recent research studies. These recent research directions should be integrated for achieving more real help. Also, a set of non-addressed research directions are suggested for future focus.

Keywords Sign language recognition · Sign language animation · Deaf-mutes · Hand gestures · Lip reading

1 Introduction

Sign language is the most expressive way for deaf and mute people. Non-deaf-mute people do not have the culture of learning sign language even if they want helping disabled people. Therefore, deaf and mute people live in isolation from other people. But, disabled people need to communicate with other people in their social life. So, sign language recognition is one of the most growing and significant research areas. Recently, many sign language recognition methods have been developed for helping deaf and dumb people [1].

Sign language is composed of some hand gestures, mouthing cues, and face expressions for representing alphabets, numbers, words, and feelings. Each country has its own sign language with different hand gestures such as American Sign Language (ASL) and Indian Sign Language (ISL). The most used sign language over the world is the ASL [2].

With the increased use of computer devices and smart phones, it became easy to exploit them to facilitate communication with deaf and mute people [3]. According to World Health Organization, deaf and mute people represent more than 5% of the world population. Therefore, it is a necessary to use technology to incorporate disabled people in communities [4]. There is a need to translate sign language to speech and speech to sign language for this communication.

In this paper, we introduce a survey for sign language translation systems and applications. A discussion of recent studies for sign language interpretation using machine learning, image processing, artificial intelligence, and animation tools is provided. Recent studies for sign language

✉ Fatma M. Najib
fatma_mohamed@cis.asu.edu.eg

1 Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

recognition using hand gestures, facial expressions translation, and lip reading are presented.

Also, recent studies for speech to sign language translation using animation are discussed. In addition, some of the popular and highly rated mobile applications that are available for Android and Apple smartphones for helping deaf and mute people in their life are presented.

In this paper, we highlight the recent research and real used applications of sign language recognition field for providing a link between research proposals and real applications. Also, we introduce a proposal of integrated functionalities coverers all separated focused directions in the recent studies. Covering these integrated research directions in real application (s) used in our real life will greatly help the deaf/ mute people in their communication. Finally in this paper, we suggest a set of non-addressed research directions for future work.

The rest of this paper is organized as follows. Section 2 illustrates sign language interpretation process. Section 3 and 4 discusses the recent studies of sign language recognition using hand gestures translation, lip reading, and facial expression interpretation. In Sect. 5, recent studies of speech to sign language conversion are presented. Section 6 discusses the real-time processing and computational requirements challenges. Section 7 presents some of the popular and highly rated mobile applications for disabled people. Section 8 presents our proposal for integrated functionalities that cover the separated addressed research directions in the recent studies for providing great real help for deaf/ mute people. Section 9 suggests set of future research directions. Section 10 is the conclusion of the paper.

Figure 1 represents an explanatory taxonomy of the sign language interpretation directions will be presented and discussed in the rest of this paper with the cited papers in each direction.

2 Sign language interpretation

There are two reverse processes for the communication between deaf-mute people and non-deaf-mute people. The sign language to speech translation process is performed by any system / software / application that help non-deaf-mute people to understand deaf and mute people. The speech to sign language process helps deaf and mute people understanding what non-deaf-mute people say.

Sign language to text/speech or text/speech to sign language translation systems is trained using dataset(s) of specific language or multiple languages involving set of signs and its corresponding text. A sign represents a letters, number, or word.

Figure 2 clarifies the sign language to speech translation process. Firstly, images or videos of sequenced images capture the signs performed by the deaf and mute people. Then, image pre-processing methods are applied such as segmentation and morphological filtering. Hand and face are continuously extracted and tracked in the translation process. Feature extraction methods are used for reduce the data dimensionality. Then, the used classifier recognizes signs. The recognized signs are translated to its corresponding text. Finally, the text is converted into speech.

Figure 3 represents the sign language to speech translation process. The input speech of non-deaf-mute people is translated into text. Speech API from Google is an example for speech to text conversion methods. Then, text pre-processing methods are used such as tokenization and lemmatization. Finally, the pre-processed text is converted

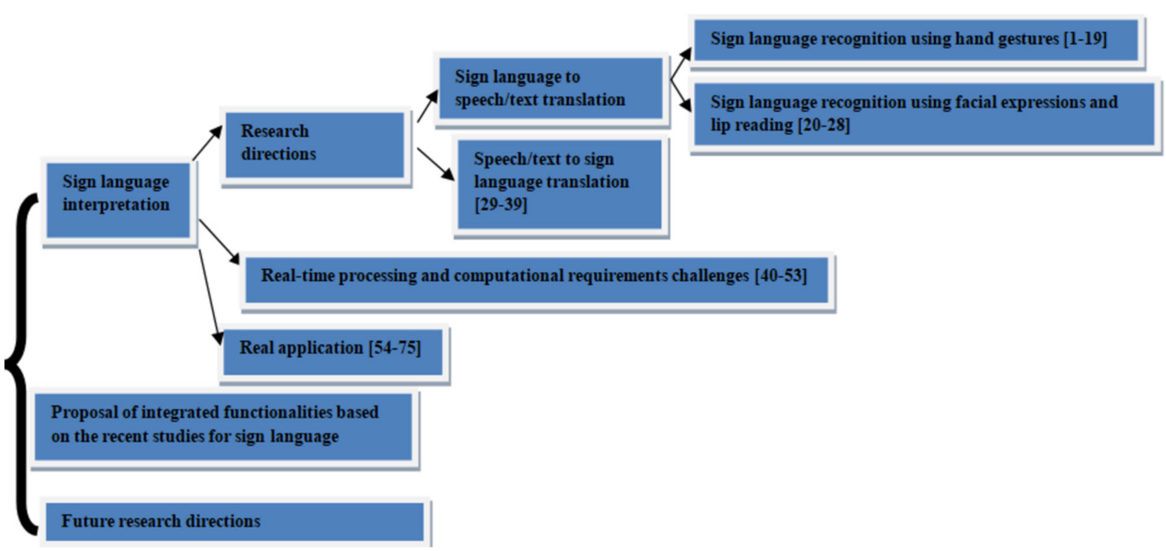

Fig. 1 Taxonomy of sign language interpretation directions in the rest of the paper

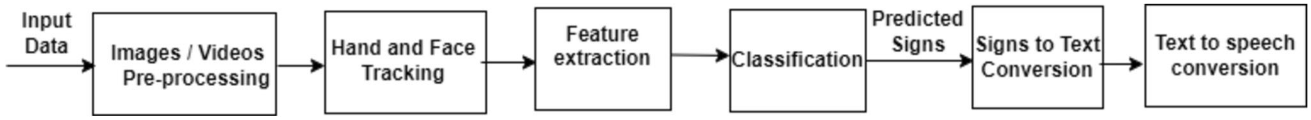

Fig. 2 Sign language to speech translation process

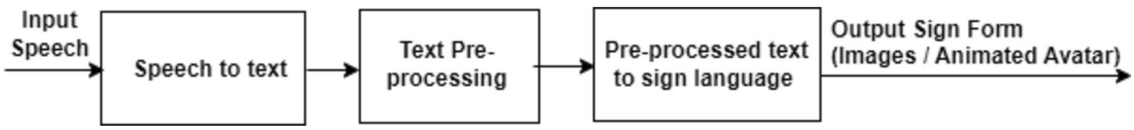

Fig. 3 Speech to sign language translation process

into its corresponding signs in the form of animated avatar or matched images.

Different sign languages are available. Each country has its own sign language. Each sign language has its own syntax and semantics. Figures 4, 5, 6, 7, 8, and 9 present

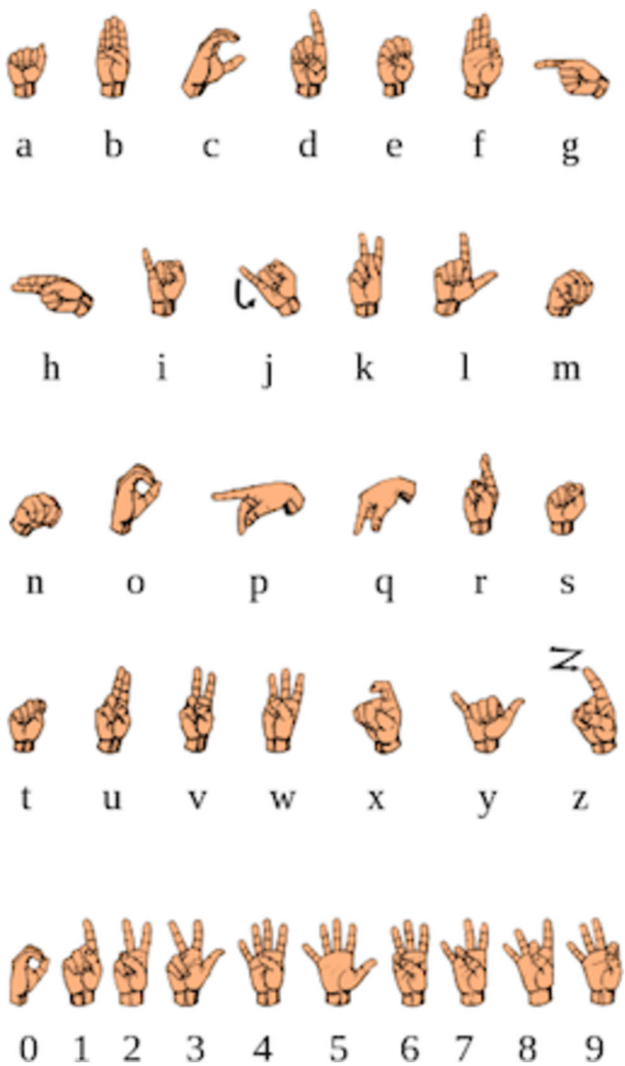

Fig. 4 American sign language alphabets

Fig. 5 British sign language alphabets

Fig. 6 Chinese sign language alphabets

Fig. 7 Indian sign language alphabets

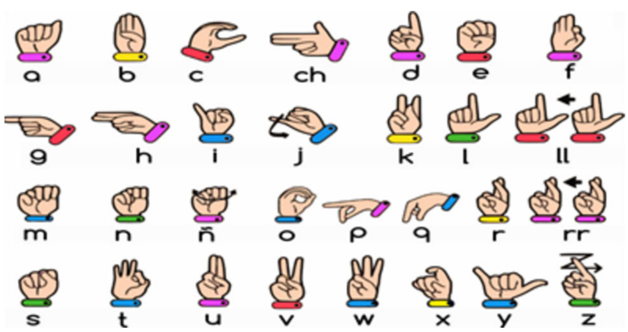

Fig. 8 Spanish sign language alphabets

alphabets of some of most popular sign languages, which are American, British, Chinese, Indian, Spanish, and Arabic Sign Languages. We can notice that each sign language has different signs from the other sign languages.

3 Sign language recognition using hand gestures

Sign language is the main communication way of deaf and dumb people. So, sign language recognition is a significant and critical issue. Most of recent studies of sign language recognition are based on hand gestures interpretation. This section discusses the recent studies for sign language recognition using hand gestures.

3.1 Recent studies of hand gestures translation

In [1], the Indian sign language was recognized using MATLAB. The Linear Discriminant Analysis (LDA) algorithm which uses linear classifier was used for recognizing hand gestures. The recognized signs were converted into text and voice format. Ten images for each sign were used in the classification process. In [2], the Convolutional

Fig. 9 Arabic sign language alphabets

Neural Network was used for sign language recognition. Images were continuously captured by the webcam with rate of 30–40 frames per second. They used OpenCV library built-in functions to process the sequenced images. The recognized signs were converted into voice with the help of Google audio built-in function. The used dataset involved signs of Letters, numbers, and static words.

In [3], they proposed a solution for recognizing hand gestures. Hand images were captured by camera and converted into the binary form. Then, OpenCV library functions were used for image processing. Hand gestures were recognized by Hand contour and convex hull detection. Finally, the recorded soundtrack of the recognized gestures was played. In [4], sign language was recognized using three few-shot machine learning algorithms which are Model-Agnostic Meta-Learning, Matching Networks, and Prototypical Networks. A small number of samples per class were used in the learning process. Recognized signs were converted into voice.

In [5], sign language recognition system was proposed using the Convolutional Neural Network (CNN). Images were trained using Keras. Training data involved images of letters, numbers, and 34 static words. The recognized signs were converted into text. In [6], deep learning-based solution was proposed for sign language recognition. Dataset of 1820 images of letters and 700 images of numbers was used in the classification process. Recognized hand gestures were converted into text and speech. In [7], a dynamic system based on Hidden Markov Models (HMM) was proposed for sign language recognition. The proposed system tracked hand movements of the user using a contour-based hand tracker. Classification of hand gestures was performed using the HMM.

In [8], they proposed a real-time Indian sign detection model based on deep neural network. The MobileNet model was used which is a CNN open sourced by Google. The recognized signs were converted into text and voice. In [9], a chat system for disabled people was proposed. It used a feed forward neural network for automatically recognition of sign language. This system got continuous web cam images that form dynamic video of sequenced hand signs. The proposed system used Hidden Markov model for converting the recognized gestures to voice.

In [10], they proposed an android mobile application to help deaf and mute people. It recognized Pakistan sign language. They used the convolutional neural network algorithm for images classification. This application recognized Pakistani sign images of letters, numbers, static words, and static sentences.

In [11], they proposed technological services that help Arabic deaf and dumb people. A smartphone chat application with Arabic signs was developed. Also, ABSHER website was proposed. It added an icon next to each field in the website and by pressing it an animation will translate this field to Arabic Sign language.

In [12], an Indian sign language conversion approach was proposed. It is based on convolutional neural network for sign language recognition. Deep learning was used for converting recognized signs to text. In [13], an intelligent approach-based Deep Learning using Gloves (IDLG) was proposed for sign language recognition. The glove contained five flex sensors and an accelerometer which provided Low-Cost Control System. IDLG was tested using different hand gestures from ten individuals.

In [14], Arabic sign language translator approach-based deep learning was proposed. They used 12 classification models which are VGG16, VGG19, ResNet50, InceptionV3, Xception, InceptionResNetV2, MobileNet, DenseNet121, DenseNet169, DenseNet201, NASNetLarge, and NASNetMobile. The majority of the 12 predictions were used. In [15], they proposed sign language recognition model-based deep learning that was built using LabelImg software and TensorFlow object detection API. For performance evaluation, 20 images for each hand sign in different angles, lighting, and backgrounds were used.

In [16], Indian sign language detection system-based deep learning was proposed. Two deep learning models were used which are LSTM and GRU feedback-based learning models. They used dataset with signs for 11 Indian words captured by 16 male and female subjects forming 1100 video samples for each word.

In [17], they proposed a deep learning-based sign language detection model. It used convolutional neural networks in the classification process. They used a dataset with 35,000 images for 100 different sign including letters, digits, and static commonly used words. In [18], a deep learning-based convolutional neural networks system for Indian sign language recognition was proposed. They used a dataset of 2500 images for Indian letters created by Rahmaniya HSS Special School Calicut.

In [19], a skeleton-aware multi-modal framework with a global ensemble model was proposed for multi-lingual sign language recognition. Datasets with three languages which are Turkish, Chinese, and American signs were used in the evaluation.

The Turkish dataset involved 226 different signs in different 20 backgrounds performed by 43 individuals. The Chinese dataset involved 500 words performed by 50 individuals. The American dataset had 2000 words performed by 119 subjects.

3.2 Discussion of recent studies for hand gestures translation

Table 1 presents a comparison between the previously mentioned studies for hand gestures translation including techniques, datasets, sign languages, and performance. Different sign languages were supported by these studies including American, Indian, Chinese, Arabic, Pakistan, and Turkish sign languages. Different machine learning and image processing techniques were used by these studies. Different datasets including hand signs for letters, numbers, and static words of different languages were used by these studies.

The skeleton-aware multi-modal framework [19] and the deep learning-based sign language detection model in [17] had the best accuracy on Turkish, Chinese, and American sign datasets in the previously mentioned hand gestures translation studies. Their accuracy was more than 99.5%. Accuracy of studies in [6, 9, 13, 14, 16, 17, 18, and 19] exceeded 95%. With regard to signs of words and sentences, they were covered by only studies in [2, 5, 7, 9, 10, 11, 14, 15, and 17]. However, the used signs were represented small static sets of words or sentences. Only [2, 7, 8, 9, 16, and 19] recognized signs from videos.

Table 1 Comparison between previously mentioned studies for hand gestures translation

References	Technique	Dataset	Sign language	Performance
[1]	The linear discriminant analysis (LDA) algorithm	Dataset of ten images for each 26 signs of Indian letters	Indian	High accuracy (The actual percentage was not mentioned)
[2]	Convolutional neural network	Datasets of Letters and numbers and set of static words	American	Accuracy = 86%
[3]	OpenCV library functions for image processing	Images of American sign language letters	American	Efficiency in the range of 75–85%
[4]	Few-shot machine learning	Images of few-shot samples of the 26 American signs	American	Accuracy approximately ranged from 60 to 90%
[5]	Convolutional neural network (CNN)	Dataset of American sign language for letters, numbers, and 34 static words	American	Average accuracy = 93.67%, Recognition time = 15 s
[6]	Deep learning	Dataset of 1820 images of letters and 700 images of numbers. Different variants of letters and numbers were shown	American	Accuracy for letters and numbers recognition = 98.46 and 98.9%
[7]	Hidden Markov models	Dataset of 1000 images. Each sign has 50 images with different positions and light levels	American	Accuracy = 60%
[8]	Deep learning	Dataset with 1000 actions. There are 30 videos for each action. Each video has a length of 30 frames	Indian	Accuracy = 92.1%
[9]	Feed forward neural network and hidden Markov model	Dataset of 1300 images of hand signs	American	Average accuracy = 96%
[10]	Convolutional neural network	Dataset obtained from kaggle involved images of Pakistani signs for letters, numbers, static words, and static sentences	Pakistan	Accuracy ranged from 80 to 90%
[11]	Web and smartphone services	Dataset contains large number of images for Arabic signs	Arabic	The actual performance was not mentioned
[12]	neural networks and deep learning	Dataset of hand signs for Indian letters	Indian	High accuracy (The actual percentage was not mentioned)
[13]	Deep learning	Dataset of 10,000 data points from 20 different people	American	Accuracy = 97%
[14]	Deep learning	Dataset of hand signs for Arabic letters and static words	Arabic	Accuracy = 93.7%
[15]	Deep learning	Dataset of hand signs of Letters and static words. 20 images were used for each sign in different situations such as different angles, lighting, and backgrounds	American	Accuracy = 85%
[16]	LSTM and GRU feedback-based learning models	Dataset contains Indian signs for 11 words. 1100 video samples were captured for each word by 16 subjects	Indian	Accuracy = 97%
[17]	Deep learning-based convolutional neural networks	Dataset of 35,000 sign images for 100 signs captured by different users	American	Accuracy = 99.72%, 99.90% for colored and grayscale images
[18]	Deep learning-based convolutional neural network	Dataset collected by Rahmaniya HSS Special School Calicut. It involved 2500 images of Indian letters gathered by seven subjects	Indian	Accuracy = 98.64%
[19]	Skeleton-aware multi-modal	Three datasets with Turkish, Chinese, and American signs	Turkish, Chinese, American	TOP-1 Accuracy of the three datasets = 99.76%, 99.95%, 84.94%

4 Sign language recognition using facial expressions and lip reading

Some studies in sign language recognition focus on hand signs translation and ignore face expressions and lip reading. But, recently there is a lot of focus on sign language recognition based on a combination of hand signs, lip reading, and facial expressions. This section presents the recent studies which were adopted perspective of combined recognition.

4.1 Recent studies of facial expressions and lip reading in sign language translation

In [20], a keyword spotting method was proposed for sign language recognition with the help of lip reading. Sign language was detected from continuous videos. Dataset with 1000 images capturing hand signs and mouthing cues was used in the evaluation.

In [21], the spatial–temporal part-aware network (StepNet) framework was proposed for sign language recognition using hand signs and facial expression. Multiple languages were supported by StepNet by using three datasets with different languages. WLASL dataset with American signs including 2000 different words was performed by 119 signers. NMFs-CSL dataset included 1,067 different words for Chinese sign language. BOBSL dataset contained 452K samples performed by 39 individuals.

In [22], they proposed Russian sign language recognition system which used lip reading in the recognition process. This method determined signer's lip region and then represents it as a feature vector of 56 features (24 for geometry features and 32 for pixel features). Face landmarks localization algorithm was used for lip leading. The proposed model was trained using 164 phrases recorded by 13 signers from the Pavlovsk boarding school for the hearing impaired.

In [23], deep learning-based sign language detection framework with the help of facial expression was proposed. A combined HOG–SVM and Haar Cascade classifiers were used for extracting facial and lip features. Then, a fine-tuned MobileNet with LSTM network was used for facial expression recognition. Two datasets with facial expression images were used in the evaluation.

In [24], they proposed automatic lip reading model for emotion recognition based on different facial expressions such as happy, sad, fear, and anger. A dataset with 7500 videos for 12 sentences recorded by 91 individuals was used in the evaluation. In [25], a hybrid model was proposed for German sign language recognition using hand signs and lip reading. Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Hidden Markov Model (HMMs) models were used in the classification process. Dataset contained 1081 different German signs performed by 9 signers was used in the evaluation.

In [26], a spatial–temporal multi-cue (STMC) network was proposed for sign language detection using hand signs and facial expressions. A 2D-CNN was used for multi-cue features generation. Recurrent neural networks (RNN) and connectionist temporal classification (CTC) were used for video sequence mapping. They used datasets for German and Chinese sign languages in the evaluation.

In [27], sign language recognition system with the help of facial expressions and hand signs was proposed using spatial–temporal map. CNN was adopted for motion recognition. OpenPose library was used for detecting the key points of hands, face, and feet of a single image. They used a dataset created by Korea Electronics Technology Institute that contained 10,480 videos with 419 words and 105 sentences recorded by ten experts. In [28], a SMPL-X parametric model was deployed for sign language recognition. It extracted 3D body shape, face, and hands information from an image. Dataset of Greek sign language was used in the evaluation.

4.2 Discussion of recent studies of facial expressions and lip reading in sign language translation

A comparison between the previously mentioned studies that considered facial expressions and lip reading in recognizing sign language is presented in Table 2. Techniques, datasets, sign languages, and performance of these studies are presented. The used datasets consist of images/videos involving different mouthing cues and facial expressions such as happy, sad, and anger in addition to hand signs. Different sign languages were supported in these studies such as American, British, Chinese, Russian, German, Greek, and Korean sign languages. Machine learning and image processing techniques are used by these studies. From all the previously mentioned studies of facial expressions and lip reading in sign language translation, studies in [21, 23, 24, and 28] achieved the highest accuracies of 98.7%, 96%, 94.04%, and 94.77% on British, American, and Greek sign datasets. Only research in [25] covered signs of alphabets, digits, words, and sentences.

5 Speech/text to sign language translation using animation

Deaf and mute people need understanding what the non-deaf-mute people say in their social communication. Many studies are concerned with this issue. This section presents

Table 2 Comparison between previously mentioned studies for facial expressions and lip reading translation

Reference	Technique	Dataset	Language	Performance
[20]	Visual keyword spotting model	Dataset with sign images for approximately 1000 words	British	Accuracy is up to 85.26%
[21]	Spatial–temporal part-aware network	Three datasets for Turkish, Chinese, and American sign language recognition	American, Chinese, British	TOP-1 Accuracy of the three datasets = 83.6%, 72.3%, 98.7%
[22]	Face landmarks localization algorithm	Dataset consists of 164 phrases recorded by 13 signers from Pavlovsk boarding school for the hearing impaired	Russian	The actual percentage was not mentioned
[23]	HOG–SVM and Haar Cascade classifiers	Miracl-VC1 dataset that contained ten words and ten phrases recorded by 15 signers	American	Accuracy = 96%
	MobileNet with LSTM network	Cohan-Kanade dataset with 2000 facial images		
[24]	3D ResNet-18 model	CREMA-D dataset that contained 7500 videos for 12 sentences recorded by 91 individuals	American	Accuracy = 94.04%
[25]	CNN, LSTM, and HMMs models	Dataset involved German sign language videos for 1081 different signs performed by 9 signers	German	Accuracy = 73.4%
[26]	Recurrent neural networks (RNN) and connectionist temporal classification	PHOENIX-2014 dataset with 1295 words recorder by 9 signers	German and Chinese	The actual percentage was not mentioned
		CSL Chinese dataset with 100 sentences including 178 words recorded by 50 signers		
[27]	Spatial–temporal map	KETI dataset that contained 10,480 sign language videos with 419 words and 105 sentences performed by ten experts	Korean	Accuracy = 84.5%
[28]	SMPL-X model	Dataset contained 3500 Greek sign language videos with 347 different signs produced by two signers	Greek	Accuracy = 94.77%

recent studies for translating speech or text to sign language using animation.

5.1 Recent studies of speech/text to sign language translation

In [29], an android mobile application was developed for translating American, Filipino text into sign language animation. It used the Knuth–Morris–Pratt Naïve Algorithm to search for the corresponding sign language animation of the input text by the user. The proposed application was tried by 44 students in Lyceum of the Philippines University Manila.

In [30], an avatar to person (ATP) online video technology was proposed. The proposed system enabled disabled people conducting face-to-face video communication. ATP converted input sign language of deaf and dumb people to facial expression animation avatar that help disabled people in online social communication and meetings. Millions of thousands samples of emotional data with different facial expressions and speaking styles were used by the training model. The system was tested by 18 participants.

In [31], intelligent gloves were proposed for facilitate the communication between deaf-mute and non-deaf-mute individuals. It sensed the hand signs and translated it into text and voice via the screen and speaker of the gloves. Also, it translated the voice of non-deaf-mute people to 3D animated avatar with sign language that displayed on the screen of the gloves.

In [32], they proposed an android mobile application that translated text to sign language using 3D virtual avatar. Urdu, Sindhi, and American sign languages were supported. Alphabets, letters, static words, and static sentences were involved.

In [33], Punjabi sign language generator system was proposed with the help of Adobe Illustrator for graphical form and Adobe Flash for 2D animation form. Punjabi letters, numbers, and some static daily words were translated to animated Punjabi sign language. They used two methods for generating animations which are tweening and step-by-step animation methods.

In [34], they proposed Indian sign language converter system that converted speech to sequence of sign language symbols. The proposed system converted speech to text using Unity 3D tool. Then, text was converted to combined phrases with Subject–object–verb patterns in the pre-processing phase. Finally, animated avatar of the corresponding translation was presented.

In [35], Gujarati sign language conversion system was proposed. The input voice was converted into Guajarati text. Then, text was converted into HamNoSys notation and SiGML XML format. Finally, animated avatar was displayed. In [36], a progressive transformer model was proposed for German sign language conversion. It converted speech into continuous 3D skeleton joint sign pose sequences with the help of dynamic time warping method. Dataset of 8257 videos with 2887 German words and 1066 different signs was used.

In [37], an Indian sign language converter was proposed. Firstly, the proposed system converted the input speech into text using Google API. Then, text was analyzed, and animated avatar with the corresponding sign language was displayed. Rule-based classification algorithm was used. They used Signing Gesture Markup Language (SiGML) to convert text into sign language.

In [38], Internet of Things-based framework for speech translation into Arabic sign language was proposed. The proposed framework consisted of three main modules which are speech recognizer, natural language processor, and videos generator. A dataset with 246 Arabic signs was used to train this system. In [39], an android mobile application was developed for converting Bangla speech into sign language animation with the help of CMU Sphinx software and Adobe character animator. A dataset with signs for more than 200 Bangla words was used.

5.2 Discussion of recent studies of speech/text to sign language translation

Table 3 presents comparison between the previously mentioned studies for speech/text to sign language translation. Techniques, datasets, sign languages, and performance of these studies are presented. Natural language processing, text pre-processing, and machine learning methods were used. Different animation tools were used. Datasets for American, Indian, German, Filipino, Urdu, Sindhi, Arabic, Punjabi, Gujarati, and Bangla sign languages were used in these studies. The best accuracy from all the previously mentioned speech/text to sign language translation studies was 90% and achieved by [38] on Arabic dataset. Only studies in [32–36, and 38] addressed signs for alphabets, digits, words, and sentences.

Also, Table 4 presents an overall comparison between the previously mentioned studies in Sect. 3, 4, and 5.

6 Real-time processing and computational requirements challenges

There are two main strategies for sign language recognition which are vision-based and sensor-based strategies [40]. The first strategy (vision based) works on camera images or video of signs. This strategy is a cost-effective, but it needs more pre-processing of the input images or video frames. Also, it requires effective cameras for high-quality images and videos. Although there is the low cost of the vision-based strategy, the recognition performance is greatly affected by the pre-processing phase and the quality of the images. Lighting low quality, lack clarity, and noisy background of the image can lead to low recognition performance. This strategy requires high pre-processing performance for enhancing the recognition accuracy [41]. Also, vision-based strategy requires high computational requirement and has more real-time computational delay [40]. One of the important advantages of vision-based strategy is taking in consideration facial expression and lip reading, but this processing requires more computational power [41, 42]. Most recent studies handle sign language recognition using hand gestures. Few studies handle facial expression and lip reading.

Computational requirements of image processing in vision-based strategy can be enhanced and reduced. Image reduction and unwanted objects removal from the original image can play role in reducing the high computational requirements. Also, converting the original image to grayscale and removing/ reducing effects of the image background can reduce computational requirements. In addition, features reduction will positively affect and reduce the computational requirements of the used classifier [41].

The second strategy (sensor based) is based on wearing intelligent gloves with sensors providing sensing information such as hand/figures shape and movement representing signs. These sensing data are captured in real time. Setup and sensor data gathering steps are performed easily. Sensor-based strategy requires lower computational power [42]. Sensor signals are more accurate and do not require the complex steps of the pre-processing phase [42]. It requires high performance of sensors to get more accurate results. Sensor-based strategy achieved higher performance than the vision-based strategy, but it requires higher cost for the expensive sensors of the gloves [40].

Lower pre-processing is required in the sensor-based strategy, but it helps enhancing the recognition performance. Removing noise from the captured raw sensor data and extracting meaningful sensing information affects the

Table 3 Comparison between previously mentioned studies for speech to animation translation

References	Techniques/tools	Dataset	Language	Performance
[29]	Knuth–Morris–Pratt Naïve Algorithm	Dataset of (words/sentences) of American, Filipino languages related to its sign language animations	American, Filipino	Overall Evaluation is Excellent with rate = 4.65
[30]	Deep neural networks and emotion regulation methods	Dataset containing millions of thousands samples with emotional data	American	24 successful translations and 12 failed translations
[31]	Flex sensors and an Arduino nanotechnologies	Database stored in the SD card of the gloves to match hand gestures with text and speech. Also, it converts speech into 3D avatar animation	American	The actual percentage was not mentioned
[32]	Polygon modeling tools	Database containing English, Urdu, and Sindhi alphabets. Also, numbers and some static words were involved	Urdu, Sindhi, American	The actual percentage was not mentioned
[33]	Adobe illustrator and adobe flash software	Dataset of Punjabi sign language involving 35 alphabets, ten numbers, and static words	Punjabi	The actual percentage was not mentioned
[34]	Natural language processing semantics, lemmatization methods, and Unity 3D tool	Dataset of Indian combined phrases with subject–object–verb patterns and their corresponding animated sign language. This dataset involved 2785 words	Indian	The actual percentage was not mentioned
[35]	NLP and CNN	Dataset contained Gujarati language signs	Gujarati	The actual percentage was not mentioned
[36]	Progressive transformer and dynamic time warping method	PHOENIX14 dataset containing 8257 videos with 2887 German words and 1066 different signs produced by 9 signers	German	BLEU and ROUGE scores up to 0.55
[37]	NLP, SVM, polarity detection method, Google API	Dataset of Indian words with its corresponding animation	Indian	Accuracy = 80%
[38]	NLP, Google API, video generator	Dataset contained signs of Arabic letters, numbers, and 200 static words	Arabic	Approximate accuracy = 90%
[39]	NLP, CMU sphinx software, adobe character animator tool	Dataset with signs for more than 200 Bangla words	Bangla	Accuracy = 88%

recognition accuracy. Sensor data may contain huge number of features. Reducing some features of the sensor data can reduce the computational requirements [42].

After the pre-processing and feature extraction phases, choosing the suitable machine learning model is a critical issue. There are many popular and extensively used classifiers in the literature such as SVM, HMM, CNN, and Deep NN. CNN and Deep NN used with large datasets and achieve highest accuracies, but require higher computational resources. SVM and HMM are suitable with small datasets and have better training and testing time than other models. CNN and Deep NN speed can be perfectly improved by enhance the processing environment such as using parallel processing [42].

6.1 Real-time sign language recognition challenge

Providing real-time gathering, processing, and recognition in the sign language recognition systems is a challenge. There is a big need for real-time, continuous, and dynamic sign language recognition to get more benefits for deaf-mutes in their daily interactions. This still represents an open challenge in recent sign language recognition systems [42]. This section presents studies used real-time methods in sign language recognition field.

In [43], a real-time deep learning-based recognition system for Assamese sign language was proposed. For providing real-time processing, they are based on MediaPipe framework that helps processing time series data. MediaPipe can apply real-time hand tracking. Images landmarks were detected using the MediaPipe environment. A deep learning technique was used in the learning process. In [44], a real-time sign recognition system was proposed based on MediaPipe Holistic model. MediaPipe Holistic model can get real-time sensing information including hand tracking and facial expression. They used deep learning eliminate background noise of sign videos.

In [45], a deep learning-based sign language recognition system for American language was proposed. They used

Table 4 Comparison between previously mentioned studies for deaf-mutes

References	Hand image	Full image (hand reading)	Video reading (hand focus)	Video reading (hand and face focus)	Conversion type 1. Sign language to text/speech 2. Text/Speech to Sign language	Signs of 1. Letters and digits 2. set of static words	Animation technology	Multi-lingual
[1]	✓				1	1		
[2]			✓		1	1 and 2		
[3]	✓				1	1		
[4]	✓				1	1		
[5]		✓			1	1 and 2		
[6]	✓				1	1		
[7]			✓		1	1 and 2		
[8]			✓		1	2		
[9]			✓		1	1 and 2		
[10]		✓			1	1 and 2		
[11]	✓				1 and 2	1 and 2	✓	
[12]	✓				1	1		
[13]	✓				1	1		
[14]		✓			1	1 and 2		
[15]		✓			1	1 and 2		
[16]			✓		1	2		
[17]		✓			1	1 and 2		
[18]	✓				1	1		
[19]			✓		1	2		✓
[20]				✓	1	2		
[21]				✓	1	2		✓
[22]				✓	1	2		
[23]				✓	1	1		
[24]				✓	1	2		
[25]				✓	1	1 and 2		
[26]				✓	1	2		✓
[27]				✓	1	2		
[28]				✓	1	2		
[29]					2	2	✓	✓
[30]					2	2	✓	
[31]			✓		2	2	✓	
[32]					2	1 and 2	✓	✓
[33]					2	1 and 2	✓	
[34]					2	1 and 2	✓	
[35]					2	1 and 2	✓	
[36]					2	1 and 2		
[37]					2	2	✓	
[38]					2	1 and 2	✓	
[39]					2	2	✓	

Jetson Nano processor for real-time detecting hand guests from captured images to be translated. In [46], a real-time virtual recognition system for American sign language was proposed based on CNN. OpenCV was used for real-time sign detection. In [47], a real-time recognition system for Indian language was proposed. They used bag of visual words, SURF (speeded up robust features), and SVM for providing real-time predictions.

A real-time recognition system for some static words of ASL was proposed in [48]. They used the MediaPipe Holistic and OpenCV for real-time capturing signs for creating their used dataset. LSTM was used in the detection process. In [49], a real-time Indian sign language translator was proposed. A real-time processing of signs was performed with the help of MediaPipe, Dynamic Time Warping, and NLP. In [50], a Turkish sign language-based skeleton-aware multi-modal and CNN recognition system was proposed. They performed parallel processing, enhanced the used hardware environment, and utilized memory for providing real-time processing.

In [51], an ASL recognition system-based CNN was proposed. OpenCV was used for real-time signs capturing to generate the used dataset. In [52], a real-time recognition system for sign language was proposed. Pre-trained models were used and fine-tuned on the used sign datasets; also trained models were converted into ONNX format for providing real-time recognition. Table 5 presents a comparison between the previously mentioned real-time sign language recognition studies.

7 Deaf and mute applications

Recently, many mobile applications support deaf and mute people and help them communicating other people in their life. This section mentions some of the popular and highly rated applications that are available for Android and Apple smartphones.

7.1 Android and Apple applications for deaf and mute people

Ntouch mobile application [53, 54] offered by Sorenson, supports Video Relay Service (VRS). Ntouch is a free application that enables making video relay calls anytime and anywhere. It offers English and Spanish sign language interpreters. Hand Talk Translator and Hand Talk [55, 56] mobile applications use artificial intelligence to generate 3D interpreter that automatically convert text and audio to sign language. Hand Talk Translator and Hand Talk applications help non-deaf or mute people to simply learn sign language. Hand Talk Translator supports American and Brazilian sign languages. Hand Talk supports American and Portuguese sign languages.

Make it big—Large text and Make it big [57, 58] mobile applications offer displaying large text on smartphones for communication with other people without speaking. These applications help communicating with deaf people. They provide Emoji support. Dark and offline modes are supported. Deaf Wake Apple application [59] is used for waking Deaf people. It provides different types of notifications or alerts for waking deaf people such as flash lighting, vibration, and/or visual alarm.

Sorenson Wavello mobile application [60, 61] helps family and friends of deaf people by enabling them to communicate with deaf people through video relay calls. ASL dictionary—sign language mobile applications [62, 63] offered by Software Studios, is considered a dictionary for ASL. It helps non-deaf / mute people to learn American sign language. It involves over 5,000 videos

Table 5 Comparison between previously mentioned studies for real-time sign language recognition

Reference	Techniques/tools	Dataset	Language	Performance
[43]	Deep learning, MediaPipe	Generated Assamese dataset consisting of 2094 images	Assamese	Accuracy = 99%
[44]	MediaPipe Holistic, Deep learning	Real-time captured signs videos	Not mentioned	Not mentioned
[45]	Deep learning	A dataset of American alphabet sign language	American	Training accuracy = 99.74%
[46]	OpenCV, CNN	An American dataset with 34,627 images for letters	American	Accuracy = 99%
[47]	bag of visual words, SURF,SVM	An Indian sign language dataset	Indian	Accuracy = 99.94%
[48]	MediaPipe Holistic, OpenCV, LSTM	An ASL dataset consisting 270 images for words signs	American	Accuracy = 100%
[49]	MediaPipe, Dynamic Time Warping, NLP	An Indian dataset for words signs	Indian	Accuracy = 93.3%
[50]	Skeleton-Aware Multi-modal, CNN	A Turkish Sign Language dataset with 3, 742 images	Turkish	Accuracy = 97.94%
[51]	OpenCV, CNN	A ASL dataset including 8000 images	American	Accuracy = 90.56%
[52]	Fine tuning, ONNX format	American and Turkish sign language datasets	American, Turkish	Accuracy = 88.57%, 99.73%

Table 6 Comparison between previously mentioned applications for deaf-mutes

Application name	Offered by	Free application	In-app purchases	Rating and reviews from 1. Google play (G) 2. Apple (A)	Downloads	Offline mode (no internet connection)	Sign language interpreter	Video Relay Service (VRS)	Text to speech/ Speech to text	Animation technology/ real signers	Languages
1. ntouch	Sorenson	✓		G. 3 + A. 4 +	500K +		✓	✓			American, Spanish
2. Hand talk translator	Hand Talk	✓	✓	G. 4.5	1M +		✓		✓	✓	Brazilian, American
3. Hand talk	Hand Talk Servicos Ltda	✓	✓	A. 4	Not mentioned		✓		✓	✓	English, Portuguese
4. Make it big—Large text for Android	Cazimir Roman	✓	✓	G. 3 +	10K +	✓					Multilanguages
5. Make it big	An Trinh	✓		A. 4.8	Not mentioned	✓					Multilanguages
6. Deaf wake	Steven Mifsud			A. 3 +	Not mentioned	✓					English
7. Sorenson Wavello	Sorenson Communications	✓		G. 3 + A. 4.4	10K +		✓	✓			English
8. ASL dictionary—sign language	Software Studios			G. 4.5 A. 4.7	10K +		✓		✓	✓	American
9. Lingvano: Sign Language—ASL	Lingvano	✓	✓	G. 4.8	500K +		✓		✓	✓	American
10. Sign video	Significan't (UK) Ltd	✓		G. 3 +	10K +	✓	✓		✓	✓	British
11. Glide—video chat messenger	Glide	✓		G. 4.2	10M +						Multilanguages
12. Glide—Live Video Messenger	Endless Technologies Ltd	✓	✓	A.4.5	Not mentioned						Multilanguages
13. Pedius	Pedius	✓		G. 3.1	10K +			✓	✓		Multilanguages
14. Pedius	Pedius srl			A. 2.6	Not mentioned			✓	✓		Multilanguages
15. Talk to deaf for android	Kickdata	✓	✓	G. 4.1	50K +				✓		Multilanguages

Table 6 (continued)

Application name	Offered by	Free application	In-app purchases	Rating and reviews from 1. Google play (G) 2. Apple (A)	Downloads	Offline mode (no internet connection)	Sign language interpreter	Video Relay Service (VRS)	Text to speech/ Speech to text	Animation technology/ real signers	Languages
16. Live transcribe and notification	Research at Google	✓		G. 4.5	1B +				✓		Multilanguages
17. Live transcribe	Zerovik Innovations Private Limited	✓	✓	A. 4.4	Not mentioned				✓		Multilanguages
18. Deaf and mute communication	Deaf Tech Corp	✓		G. 3 +	10K +	✓			✓		Multilanguage
19.Deaf communication (Pro)	Deaf Tech Corp			G. 3 +	100 +	✓			✓		Multilanguages

containing more than 5000 words interpreted into ASL. Each sign is represented by 460 multiple ways. Also, Lingvano: Sign Language–ASL [64] mobile application provides easy and fast learning of ASL for beginners. It involved sign videos performed by deaf people.

SignVideo [65] mobile application provides British Sign Language interpreter and provides Video Relay Services. Users of this application can call any SignDirectory through a SignVideo BSL interpreter and benefits from more than 150 services including banks, insurance, and health centers. Glide—Video Chat Messenger and Glide—Live Video Messenger mobile applications [66, 67] provide fast video messaging and allow group chatting and uploading videos to social media.

Pedius mobile application [68] allows phone calls for the deaf people. Talk to Deaf for Android mobile application [69] facilitates communication with deaf and non-deaf people. This application can translate voice input of non-deaf person to text to be read by deaf person.

Live Transcribe & Notification mobile application [70] helps deaf and mute people by real-time translation of speech to text. It involved more than 80 languages. Also this application provides flashing light or vibration alarms for deaf and mute people in the risky situations such as siren and baby sounds. Live Transcribe apple mobile application [71] translates speech to text and supports more than 60 languages.

Deaf & Mute Communication free android mobile application that offered by Deaf Tech Corp is available to help disabled people communication [72]. It provided speech to text and text to speech conversion. It supports 140 languages and can work in offline mode without internet connection. Also, Deaf Tech Corp offers Deaf Communication (Pro) android mobile application for deaf people [73]. This is a paid application with no In-advertisements. It converts speech to text or text to speech with 140 languages in offline mode.

7.2 Discussion of Android and Apple applications for deaf and mute people

Table 6 presents a comparison between the previously mentioned applications for deaf and mute people. These applications are available in Google store and Apple store of Android and Apple smartphones. Most of them are free applications, but In-App Purchases are required by some of these applications. Most of applications are multi-lingual applications. Millions and thousands of people whether they are deaf-mutes or non-deaf-mutes downloaded these applications.

Most of them require internet connection. Text to speech or speech to text translation is provided by most of these applications. Some of them support Video Relay Service

(VRS). Non-deaf-mute people can use some of these applications to learn few hand signs helping them communication deaf/mute people. Sign language interpreter and animation technology are supported by some of these applications.

8 Proposal of integrated functionalities based on the recent studies for sign language recognition

Based on the mentioned studies in Sect. 3,4, and 5 for sign language recognition, we introduced a proposal of integrated functionalities that were separately covered in the recent studies. Some of the previously mentioned studies addressed sign language conversion to text/speech, but others addressed text/speech conversion to sign language. The used signs were diversified in the mentioned studies involving letters, digits, some few static words, or few static sentences.

The signs were gotten/captured using images or videos. The used images were hand signs or full images. Some mentioned researches only focused on the hand sign in the full images. Other studies focused on hand sign, facial expressions, and lip reading of the full images. With regard to studies addressed sign videos, some of them focused only the hand sign in videos, but others focused also on facial expressions and lip reading. Animation/avatar technology was used by some studies.

So, we introduce a proposal of a sign language recognition system/application that integrates all the separate handled functionalities by the recent studies. The functionalities of the proposed integrated system/application are summarized as follows:

1. Sign language to text/speech conversion.
2. Text/speech to sign language conversion.
3. Signs for letters, digits, words, and sentences are involved.
4. Covering multiple sign languages.
5. Images and videos sign options are involved.
6. Hand sign, facial expressions, and lip reading are involved in the recognition process.
7. Animation/avatar technology is involved

9 Future research directions

Most of recent studies of sign language recognition used signs of letters and digits. Static small set of words or sentences signs were used by few studies.

There are no existing studies addressed sign language recognition system with many words and sentences based on language dictionaries.

Also, online services for deaf and mutes did not get attention in the recent researches. There are no recent studies proposed solutions to facilitate the online services to dead mute people. For example, there are no studies focused on introducing solutions for the online shopping, the bank transactions, and other online services for deaf-mute people.

In addition, there is no focus on introducing social networks for dealing with deaf-mutes. Introducing research solutions for providing online social societies that are suitable to deaf-mutes is very important non-addressed issue.

There are no existing proposals for extra self-education and reading online articles, books, and other documents for deaf-mutes.

10 Conclusion

According to World Health Organization, deaf-mutes became more than 5% of the world population. There are difficulties in deaf-mutes communication with non-deaf-mutes who do not understand sign language. In real life, most of non-deaf-mutes do not understand sign language. Therefore, sign language recognition and interpretation are considered one of the challenging research areas in recent years. We study state of the art in sign language recognition and interpretation. Most of recent research utilizes machine learning, image processing, artificial intelligence, text pre-processing, and animation tools for sign language to text/speech and text/speech to sign language translation.

Sign language conversion processes which are: sign language to text/speech and text/speech to sign language are illustrated at first. Recent studies in the period from 2018 to 2023 are discussed, including (1) Sign language recognition studies using hand gestures, (2) Sign language recognition studies using combined hand gestures translation, lip reading, and facial expression interpretation, (3) Speech or text to sign language conversion studies using animation. Techniques, datasets, sign languages, and performance of these studies are briefly presented. An overall comparison of all these studies is introduced. Also, some of popular and highly rated Android and Apple mobile applications for deaf-mutes communication are discussed.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Data availability The data that support the findings of this study are available from studies and links of references section.

Declarations

Conflict of interest The author has declared no conflicts of interest for this article.

References

1. Mk NB (2018) Conversion of sign language into text. Int J Appl Eng Res 13(9):7154–7161
2. Mishra SK, Sinha S, Sinha S, Bilgaiyan S (2019) Recognition of hand gestures and conversion of voice for betterment of deaf and mute people. In Advances in computing and data sciences: third international conference, ICACDS 2019, Ghaziabad, India, April 12–13, 2019, Revised Selected Papers, Part II 3, pp 46–57. Springer Singapore
3. Manikandan K, Patidar A, Walia P, Roy AB (2018) Hand gesture detection and conversion to speech and text. arXiv preprint arXiv:1811.11997.
4. Shovkoplias G, Tkachenko M, Asadulaev A, Alekseeva O, Dobrenko N, Kazantsev D, Vatian A, Shalyto A, Gusarova N. Support for communication with deaf and dumb patients via few-shot machine learning
5. Tolentino LKS, Juan RS, Thio-ac AC, Pamahoy MAB, Forteza JRR, Garcia XJO (2019) Static sign language recognition using deep learning. Int J Mach Learn Comput 9(6):821–827
6. Saleem MI, Siddiqui A, Noor S, Luque-Nieto MA, Otero P (2022) A novel machine learning based two-way communication system for deaf and mute. Appl Sci 13(1):453
7. Saleh BM, Al-Beshr RI, Tariq MU (2020) D-talk: sign language recognition system for people with disability using machine learning and image processing. Int J Adv Trends Comput Sci Eng 9(4)
8. Sivakumar P, Amrithaa IS, Sandhiya A, Janani T, Karthikeyani S (2022) Translating Indian sign language to text using deep learning. Int J Innov Sci Res Technol 7(5):1285–1290
9. Pandey A, Chauhan A, Gupta A (2023) Voice based sign language detection for dumb people communication using machine learning. J Pharm Negative Res, pp 22–30
10. Khaskheli AH, Mirani SH, Arain A. An android app for deaf and dumb people (Sign language Recognition–Smart Talk App)
11. Alhassan F, Alqanny A, Elharbi M, Elharbi M (2020) Technology-based services for deaf and dumb people. Int J Data Sci 5(2):160–167
12. Srinivasan N, Akash, B, Aathish S, Ravesh Raaju P. Gesture2Text: a novel approach for real-time sign language conversion
13. Elwahsh H, Elkhouly A, Nasr EA, Kamrani AK, El-Shafeiy E (2022) A new intelligent approach for deaf/dumb people based on deep learning. Comput Mater Contin 72:6045–6060
14. Nahar KM, Almomani A, Shatnawi N, Alauthman M (2023). A robust model for translating arabic sign language into spoken arabic using deep learning. Intell Automat Soft Comput 37(2)
15. Sanmitra PR, Sowmya VS, Lalithanjana K (2021) Machine learning based real time sign language detection. Int J Res Eng Sci Manage 4(6):137–141
16. Kothadiya D, Bhatt C, Sapariya K, Patel K, Gil-González AB, Corchado JM (2022) Deepsign: sign language detection and recognition using deep learning. Electronics 11(11):1780
17. Wadhawan A, Kumar P (2020) Deep learning-based sign language recognition system for static signs. Neural Comput Appl 32:7957–7968
18. Sruthi CJ, Lijiya A (2019) Signet: A deep learning based indian sign language recognition system. In: 2019 International conference on communication and signal processing (ICCSP), pp 0596–0600, IEEE
19. Jiang S, Sun B, Wang L, Bai Y, Li K, Fu Y (2021) Sign language recognition via skeleton-aware multi-model ensemble. arXiv preprint arXiv:2110.06161.
20. Albanie S, Varol G, Momeni L, Afouras T, Chung JS, Fox N, Zisserman A (2020) BSL-1K: Scaling up co-articulated sign language recognition using mouthing cues. In: Computer vision–ECCV 2020: 16th European conference, Glasgow, UK, 23–28 Aug 2020, Proceedings, Part XI 16, pp 35–53, Springer International Publishing
21. Shen X, Zheng Z, Yang Y (2022) Stepnet: spatial-temporal part-aware network for sign language recognition. arXiv preprint arXiv:2212.12857
22. Ivanko D, Ryumin D, Karpov A (2019) Automatic lip-reading of hearing impaired people. Int Arch Photogramm Remote Sens Spat Inf Sci 42:97–101
23. Kumar GA, William JH (2022) Development of visual-only speech recognition system for mute people. Circuits Syst Signal Process, pp 1–21
24. Ryumina E, Ivanko D (2022) Emotional speech recognition based on lip-reading. In: International conference on speech and computer, pp 616–625. Springer International Publishing, Cham
25. 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
26. Zhou H, Zhou W, Zhou Y, Li H (2020) Spatial-temporal multi-cue network for continuous sign language recognition. Proc AAAI Conf Artif Intell 34(7):13009–13016
27. Choi SG, Park Y, Sohn CB (2022) Dataset transformation system for sign language recognition based on image classification network. Appl Sci 12(19):10075
28. Kratimenos A, Pavlakos G, Maragos P (2021) Independent sign language recognition with 3d body, hands, and face reconstruction. In: ICASSP 2021–2021 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 4270–4274, IEEE
29. Masbate AMM, Ong MBG, San Juan FRT, Paglicawan AKE, Caballero AR (2020) Text to english and text to filipino sign language translator: an android-based mobile application with text analysis using KNUTH-morris-Pratt naive algorithm. Elem Educ Online 19(4):78–84
30. Guo Z, Wang Z, Jin X (2021) "Avatar to Person"(ATP) virtual human social ability enhanced system for disabled people. Wirel Commun Mob Comput 2021:1–10
31. Babour A, Bitar H, Alzamzami O, Alahmadi D, Barsheed A, Alghamdi A, Almshjary H (2023) Intelligent gloves: an IT intervention for deaf-mute people. J Intell Syst 32(1):20220076

32. Bhatti Z, Muhammad F, Malik HAM, Hussain M, Chandio H, Channa S, Mahar Z (2021) Text to animation for sign language of Urdu and Sindhi. IKSP J Emerg Trends Basic Appl Sci 1(1):08–14
33. Kaur R, Kaswan S (2022) Conversion of Punjabi sign language using animation. In: Rising threats in expert applications and solutions: proceedings of FICR-TEAS 2022, pp 175–185, Springer Nature Singapore, Singapore
34. Sumana M, Hegde SS, Wadawadagi SN, Sujana DV, Narasimhan VG (2022) Smart tutoring system for the specially challenged children. In: Society 5.0: smart future towards enhancing the quality of society, pp 113–130. Springer Nature Singapore, Singapore
35. Aasofwala N, Verma S, Patel K (2021) A novel speech to sign communication model for gujarati language. In: 2021 Third international conference on inventive research in computing applications (ICIRCA), pp 1–5, IEEE
36. Saunders B, Camgoz NC, Bowden R (2020) Progressive transformers for end-to-end sign language production. In: Computer vision–ECCV 2020: 16th European conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI 16, pp 687–705, Springer International Publishing
37. Amal H, Reny RA, Prathap BR (2020) Hand kinesics in indian sign language using NLP techniques with SVM based polarity. Int J Eng Adv Technol IJEAT 9(4):2249–8958
38. Mouti S, Rihawi S (2020) IoT and sign language system (SLS). Int J Eng Res Technol 13(12):4199–4205
39. Kamrul AB, Sharmin, S (2020) An android communication platform between hearing impaired and general people 12(27):2249–0868
40. Alhafdee AH, Abbas H, Shahadi HI (2022) Sign language recognition and hand gestures review. Kerbala J Eng Sci 2(4):209–234
41. Madhiarasan DM, Roy P, Pratim P (2022) A comprehensive review of sign language recognition: Different types, modalities, and datasets. arXiv preprint arXiv:2204.03328. 2022 Apr 7
42. Papatsimouli M, Sarigiannidis P, Fragulis GF (2023) A survey of advancements in real-time sign language translators: integration with IoT technology. Technologies 11(4):83
43. Bora J, Dehingia S, Boruah A, Chetia AA, Gogoi D (2023) Real-time assamese sign language recognition using mediapipe and deep learning. Proc Comput Sci 1(218):1384–1393
44. Zhao Y, Zhang B, Ma K (2022) RMSLRS: Real-time multi-terminal sign language recognition system. In: International conference on innovations in bio-inspired computing and applications 2022 Dec 15 (pp 575–585). Springer Nature Switzerland, Cham
45. Yashmita BS, Chaudhary A, Kaur B, Reddy SR, Anand R (2023) Unlocking the power of Al: A real-time translation of sign language to text. In: International conference on artificial intelligence of things 2023 Mar 30 (pp. 314–330). Springer Nature Switzerland, Cham
46. Zizoune A, Zizoune A, Hamdaoui R, Salaheddine K, Riadsolh A, Ziti S (2023) Real-time implementation of an AI-based virtual sign language recognition and interpretation system. EasyChair 13
47. Kurre T, Katta T, Burla SA, Neelima N (2023) Real-time Indian sign language recognition using Image Fusion. InInternational Conference on Communications and Cyber Physical Engineering 2018 (pp. 599–605). Springer Nature Singapore, Singapore
48. Gaikwad R, Admuthe L (2022) Real-time sign language recognition of words and sentence generation using mediapipe and LSTM. In: International joint conference on advances in computational intelligence (pp. 287–297). Springer Nature Singapore, Singapore
49. Mankar Y, Wankhede Y, Thune V, Khawshe S, Sahu M. Real-Time Sign Language Interpreter Using MediaPipe, Dynamic Time Warping, and NLP. InInternational Conference on Multi-Strategy Learning Environment 2024 Jan 12 (pp. 87–98). Singapore: Springer Nature Singapore.
50. Hori N, Yamamoto M. Real-Time Isolated Sign Language Recognition. InInternational Conference on Frontiers of Artificial Intelligence, Ethics, and Multidisciplinary Applications 2023 Sep 25 (pp. 445–458). Singapore: Springer Nature Singapore.
51. Meka PK, Parusu YR, Vaddi R. Real-Time Sign Language Detection Using OpenCV. InInventive Systems and Control: Proceedings of ICISC 2023 2023 Jun 15 (pp. 523–534). Singapore: Springer Nature Singapore.
52. Novopoltsev M, Verkhovtsev L, Murtazin R, Milevich D, Zemtsova I (2023) Fine-tuning of sign language recognition models: a technical report. arXiv preprint arXiv:2302.07693. 2023 Feb 15.
53. Ntouch. https://play.google.com/store/apps/details?id=com.sorenson.mvrs.android&hl=en_US. Accessed Aug 2023
54. Ntouch. https://apps.apple.com/us/app/ntouch/id441554954. Accessed Aug 2023
55. Hand Talk Translator. https://play.google.com/store/apps/details?id=br.com.handtalk&hl=en_US. Accessed Aug 2023
56. Hand Talk. https://apps.apple.com/us/app/hand-talk/id659816995. Accessed Aug 2023
57. Make it big - Large text, https://play.google.com/store/apps/details?id=com.suspedeal.makeitbig&hl=en_US. Accessed Aug 2023
58. Make It Big. https://apps.apple.com/us/app/make-it-big/id479282584. Accessed Aug 2023
59. Deaf Wake. https://apps.apple.com/au/app/deaf-wake/id1217445602?l=sv. Accessed Aug 2023
60. Sorenson Wavello. https://play.google.com/store/apps/details?id=com.sorenson.dhv&hl=en_US. Accessed Aug 2023
61. Sorenson Wavello. https://apps.apple.com/us/app/sorenson-wavello/id1499397260. Accessed Aug 2023
62. ASL Dictionary - Sign Language. https://play.google.com/store/apps/details?id=com.softwarestudios.asldictionarysignlanguage&hl=en_US. Accessed Aug 2023
63. ASL Dictionary Sign Language. https://apps.apple.com/us/app/asl-dictionary-sign-language/id682472857. Accessed Aug 2023
64. Lingvano: Sign Language - ASL. https://play.google.com/store/apps/details?id=com.lingvano.app&hl=en_US. Accessed Aug 2023
65. SignVideo. https://play.google.com/store/apps/details?id=me.signvideo.android&hl=en_US. Accessed Aug 2023
66. Glide - Video Chat Messenger. https://play.google.com/store/apps/details?id=com.glidetalk.glideapp&hl=en_US. Accessed Aug 2023
67. Glide—Live Video Messenger. https://apps.apple.com/us/app/glide-live-video-messenger/id588199307. Accessed Aug 2023
68. Pedius. https://play.google.com/store/apps/details?id=com.client.pedius&hl=en_US. Accessed Aug 2023
69. Talk to Deaf for Android. https://play.google.com/store/apps/details?id=unique2040.com.text2speech&hl=en_US. Accessed Aug 2023
70. Live Transcribe & Notification. https://play.google.com/store/apps/details?id=com.google.audio.hearing.visualization.accessibility.scribe&hl=en_US. Accessed Aug 2023
71. Live Transcribe. https://apps.apple.com/us/app/live-transcribe/id1618181875. Accessed Aug 2023
72. Deaf & Mute Communication. https://play.google.com/store/apps/details?id=com.pack1.deafcommunication&hl=en_US. Accessed Aug 2023
73. Deaf Communication (Pro). https://play.google.com/store/apps/details?id=com.pro.deafcommunication. Accessed Aug 2023

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.