JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SIGN LANGUAGE RECOGNITION SYSTEM: REVIEW STUDY

[1]Pallavi Shantaram Rane, [2]Sangita D. Bharkad
[1]M.tech Student, [2]Assistant Professor
[1]Electronics Engineering,
[1]Government Engineering collage (Autonomous), Aurangabad (MH), India

Abstract: - Communication is the process of meaningful interactions among human beings. Natural language channels such as speech, writing, and body language (gestures) such as hand movements, head gestures, face expression, lip motion, and so on are methods for humans to connect. Understanding sign language is just as important as understanding normal language since sign language is the best way to communicate with the deaf and blind people. Voluntary movements of hands and fingers are used in sign language to express clear action. There is no universal sign language, just as there is no universal spoken language, because each nation has its own dialect of sign language. As a result, Indian sign language is utilised in India. Because sign language is so distinct from natural language, interpreters of sign language are necessary to convert common language into sign language and vice versa. As such specialists are rare in India in comparison to deaf and deaf individuals, a sign language recognition system is necessary to bridge the communication gap between the hearing-impaired cum speech-impaired population and the rest of society by translating Indian sign language into regular text. There are several approaches to sign language recognition systems that have been developed during the last decade, but there are still certain issues that need to be addressed. In existing systems, various machine learning methods are utilised.

IndexTerms: SLR(Sign Language Recognition), ISL(Indian Sign Language), CNN(Convolutional Neural Networks), PNN(Probabilistic Neural Networks), KNN(K-Nearest Neighbor), SVM(Support Vector Machines), RNN(Recurrent Neural Network), HOG(Histogram Oriented Gradient), HMM(Hidden Markov Model).

1. Introduction

Communication is the mechanism by which humans engage in meaningful ways. We interact with each other either by listening or talking in our day today life, but we find it is difficult to communicate with deaf and dumb people in our society as we are unable to understand their language and vice-versa. Sign language is a form of interaction where each word or alphabet is associated with a different sign. Rather of sound, it transmits signals through hand shapes, direction and action of the hands, arms, or body, face gestures, and lip motions. The spoken language is not universal. American sign language, Indian sign language, Chinese sign language, British sign language and French sign language are some of the most frequently used sign languages in the universe. There are almost 18 million people are deaf, around 1 percent of total Indian population according to India's National Association of Deaf. You cannot communicate with ordinary folks. Translators who can convert sign language to speech and vice versa are required to overcome the communication gap. However, there is a limited availability, and it is expensive, for the creation of the sign language recognition system that can translate signs to text or voice automatically. The Indian sign language is adopted in India. The Indian sign language is more complicated, contains both static and dynamic motions and complex movements of the hands, single and double hand.

Figure 1 depicts the representation of alphabetical gestures in Indian sign language, in which double handed notations are used for all alphabets excluding C, I, J, L, O, U, V, W and single handed sign notations are used for the remaining alphabets. Mostly double handed gestures are used for words and phrases. Some words use movements of hands like crawl, pledge, and, all etc. Facial Expressions are also used in some words like Good, no, bed, wait, dinner etc. Sentence creation is accomplished through the use of both single and double handed sign language motions. After recognition, Sign can be converted into speech or text.

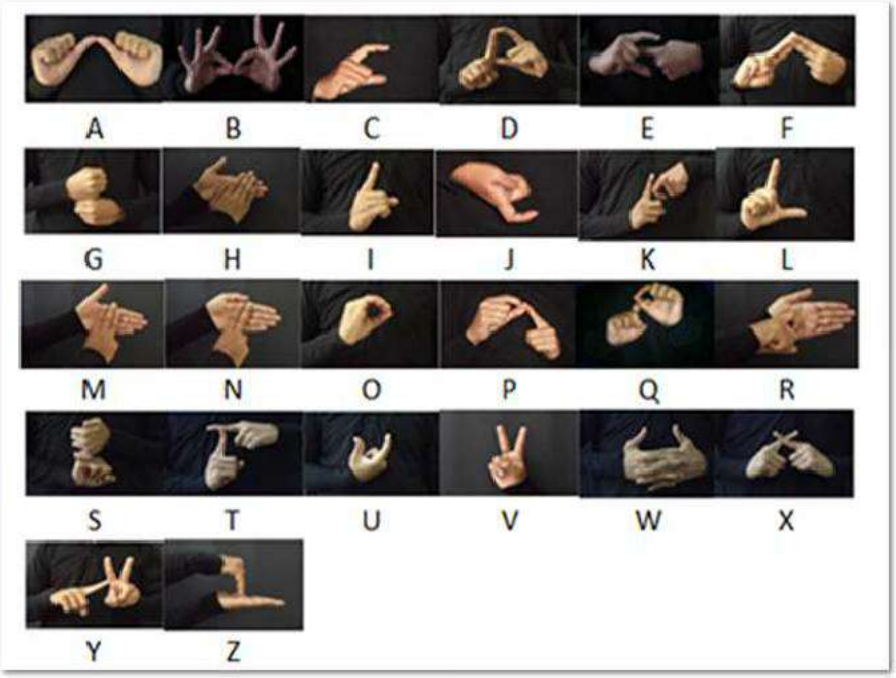

Fig.1 Alphabetical Gestures in Indian Sign Language.

2.Different approaches for sign language recognition.

Sensor-based approach and vision-based approach are two primary techniques utilised in sign language recognition. Sensor-based approach, also known as Glove-based technique. Combined approach using both systems also can be used**.** Sensor based approach uses wireless [1] or wired sensor. In this approach, signer needs to wear glove and is used for capturing the gestures. It uses some sensors like the proximity sensor, accelerometer sensor, abduction sensor and flexion sensor to extract features describing the hand sign. During processing, this technique improves the work of segmentation. The disadvantage of this method is that while the system is in operation, the signer must wear both the sensor equipment and the gloves. Using machine learning techniques, an integrated camera, and digital image processing algorithms, the vision-based approach extracts and recognises hand signals as well as the signer's facial expressions. This approach does not require any new hardware, however there are certain accuracy concerns with image processing techniques that have yet to be addressed. 3D model based and appearance based are the two approaches to vision-based sign language recognition. Using 3D information on important elements of the body parts, 3D model-based approaches extract numerous significant properties such as palm position, joint angles, and so on. This technique employs volumetric, skeletal, or a mix of the two models. This method is quite computationally costly, and systems for live analysis are still in the works. Images or videos are used as inputs in Appearance-based systems. These videos/images are utilised for interpretation directly. The body is not represented spatially. Using a template database, the parameters are extracted straight from the images/videos. Some of the templates are deformable two-dimensional representations of human body parts, notably hands. Deformable templates are approximation points in object outline approximation, which is a set of points on the object's outline. Linear interpolation functions are the most basic. These template-based models may be used for more than only hand tracking; they can also be used to classify simple gestures. Appearance-based models, which employ picture sequences as gesture templates, are a second technique in sign language gesture recognition.

3. Related work

3.1 Recognition of alphabets and numericals

The framework consists of Indian sign language recognition by fingerspelling [1]. To recognise motions, this system employs digital image processing methods and artificial neural networks. In first step gestures are captured and pre-processed. Self-created data set is used. In second stage of segmentation hand is extracted from captured images, after that feature extraction is applied in third step. In last training and testing phase ANN is trained to classify 36 signs with 360 images in training dataset and in testing phase 180 images used to test the system. The average recognition rate after experimentation is 91.11%. Another Framework [2] works on this concept is an app "Fingerspelling" which is made using open-source tool such as Blender consisting of three basic modules. These three modules are 1] "Learn" module 2] "Practice" module 3] "Test" module. Preliminary Testing is done at a hearing and speech impaired school. There were 78% of the students inspired and they said that this app can enhance the self-learning skill of ISL. Vision-based strategy in which a web camera is used to record images, which are then analysed using MATLAB and feature Extraction and classification is performed on the gesture. Hand gestures that are captured, then converted into speech and text. English and Hindi languages are used. 82% accuracy is achieved [3]. Deep Learning Approach in which Convolutional Neural Networks are employed for classification. In first phase, classifier model is made using the numeral sign and keras implementation of CNN in python platform. In the second stage, skin

segmentation is performed to determine the region of interest in the frame using another real-time system, and then the sign is predicted as it is input into the classifier model. Accuracy for the same subject is found to be 99.56 % and 97.26% for the low light conditions [4]. Another framework works on the recognition of Indian Sign Language Alphabets. Categorization is done either single-handedly or double-handedly in this case. A series of training images is extracted using feature extraction techniques such as HOG and SIFT and then concatenated in a single matrix form. The test images are split into single and double-handed motions, and classification is performed using fusion of HOG and SIFT descriptor fusion with K-nearest correlated neighbour. The trainings section contains 520 images and the test half contains 260 images. When single and double handed gestures are not classified, the accuracy rate using SIFT is 78.84%, while the accuracy rate using HOG is 80%. Accuracy rate is 90 % when both the techniques are fused. SIFT obtains an accuracy rate of 92.5 % for single handed motions and 75.55 % for double handed motions, when single and double-handed motions are classified. HOG achieves a single handed accuracy percentage of 100% and a double handed accuracy rate of 82.77%. The combination of the two techniques results in a 97.5 % single handed and a 91.11 % double handed result [5]. Another study on the recognition of Indian Sign Language in which dataset is divided into training and testing. After that image are pre-processed and then converted into binary format. The HOG features and the geometric features are calculated for the image. For classification, SVM and KNN are utilised. SVM improves KNN in terms of accuracy, according to experimental results. SVM in combination of HOG feature can provide an accuracy of 94.23% [6]. Automatic recognition system [7] of Indian sign language in complex backgrounds in which Computer vision, natural language processing, and pattern recognition are used. Webcam or mobile camera is used for the purpose of image acquisition. Preprocessing and skin segmentation is done by dividing video taken into frames. Feature extraction is carried out, and the gesture classification method is used to recognise the sign. SVM is used for classification. The work is completed in python platform and compared with cv2 library. After that, the gesture is predicted and converted into the text format.

3.2 Recognition of Words and phrases

In framework of Argentinean sign language gestures, 2300 videos are being used as a dataset for different words belonging to 46 gesture categories and 50 videos per gestures are used for better recognition. To train these models, two learning methodologies are used. To train these models, two learning methodologies are used. First is spatial features CNN to train model, and second is temporal features RNN. Accuracy is 95.217% [8]. A framework proposed a sign language recognition utilising a leap motion controller and Kinect sensors, in which the sign language is recorded by leap motion and the signers' facial data is captured by a Kinect camera. Dynamic gestures datasets are used. For recognition, HMM (Hidden Markov Model) is used and after that independent Bayseian classifier approach is being applied on them. In single and double hand gestures, accuracy is 96.05% and 94.27%, respectively [9]. Indian Sign Language gestures recognition algorithm in which the numbers, alphabets, and a few words are translated into English. This algorithm performs data capture and gesture pre-processing, which is then employed by a combinational algorithm to track hand movement, and template matching is used for recognition. A self-created database with 130000 videos is used. The system's database was generated using 72000 videos, and the system's performance was tested using 58000 videos. The obtained accuracy is 97.50% [10]. The tracking techniques like the movement, shape, direction are compared for ISL [11] and analyzed by following some steps like data acquisition from the self-created database then Preprocessing for frames which will enhance the quality of images and focus on the gesture part. Meanshift and Kalman Filter are used. After that, classification and Recognition is done. ANN, SMO, Naïve Bayes classifiers and multilevel perceptron are used. The rate of recognition is equal to the correctly identified gesture videos to the total hand gesture videos. The results shows that meanshift gives more Elapse time, Precision on, and Effect of Velocity change whereas kalman filter gives more effect of illumination. Automated sign interpreter [12] that transforms Indian sign language into audio. A sensor-based approach, such as an instrumented glove, is used in this system. Some hardware component like Arduino uno, flex sensors, gyroscope with accelerometer, Bluetooth and SD-Card Module are used in this architecture. During experimentation, 90 words are tested out of 3000 words.

3.3 Recognition of Sentence

A framework [13] which focuses on invariant backgrounds for dynamic gesture recognition of ISL. The frame overlapping method was used to track relevant motions. After those frames are extracted which has maximum information for speedup. Then, for extracting features, the Discrete Wavelet Transform (DWT) is used, and for testing probing gestures, the HMM is used, which has a low time complexity and space complexity. Another framework [14] translates the sign language in written text format and it also convert it into voice form. It also provides the facility for the normal people to enter the input as a text and then system converts it into sign language images and in the video form. This system consists of three modules training, testing, and text to speech engine. It is implemented using MATLAB with PNN KNN classifier. Accuracy obtained for image is 87.04% and 90.74% for KNN and PNN respectively. Accuracy obtained for video is 86.67% and 93.33% for KNN and PNN respectively. A real-time system that has high accuracy rates and can recognise gestures in ISL. The gestures are captured by the Android smartphone, and the frames are then sent to the server. Preprocessing is performed with the help of face detection and elimination using HOG features and SVM classifier. Then skin segmentation is performed using YUV+RGB Color Space. Following skin colour segmentation, morphological operations such as erosion and dilution are utilised to reduce any noise that has been created. Furthermore, a grid-based feature extraction technique is applied, and finally, KNN and HMM classifiers are used for

classification [15]. A framework on Indian Sign Language for mobile platforms [16]. Using the front camera of a smartphone in selfie mode, a video-based dataset for Indian sign language is developed. To generate sign language feature space, video frames are pre-filtered, segmented, and feature extracted. Signs are converted to text or voice using the minimal distance classification and sign language feature space. The accuracy is approximately 90.58%.

Table 1. Different Techniques with Their Accuracy

Author Name	Method	Accuracy
Umang Patel	PNN and KNN, Contour Based and region based shape representation and description methods	82%
Sajanraj T D	CNN, Le Net Architecture.	99.56% and in low light 97.26%
Bhumika Gupta	KNN, HOG, SIFT	For Single handed 97.50% double handed 91.11%
Sarfaraz Masood	CNN, RNN, Prediction and Pool Layer Approach	95.20%
Pradeep Kumar	DMF algorithm, Kinect sensor, Leap motion sensor, HMM, feature vector, IBCC, PCA Approach	For Single handed 96.05% double handed 94.27%
Purva C. Badhe	Fourier Descriptors, Canny Edge Detection	97.50%
Shwetha S kulloli	KNN,PNN	For image in KNN 87.04% and PNN 90.74% & for video in KNN 86.67% and PNN 93.33%

4 System Overview

sign language recognition system was designed by various developers for a range of applications, with different stages of recognition. They all agree on the general architecture of the recognition system. The four stages involved in a sign language recognition system are image capture, pre-processing and segmentation, feature extraction and classification, as illustrated in fig.4.

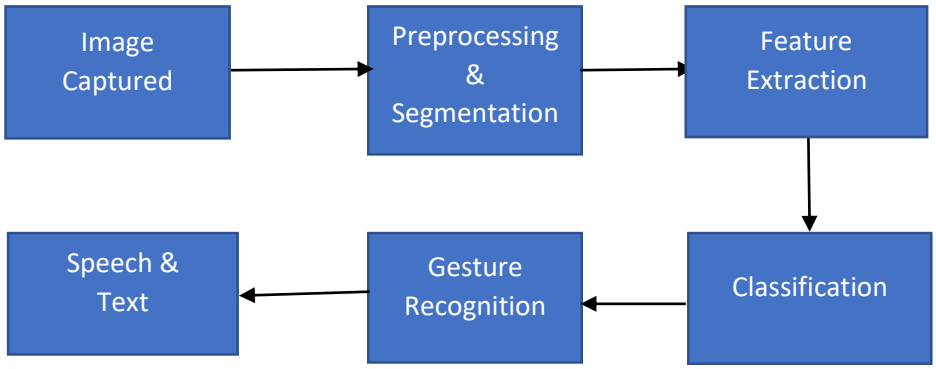

Fig. 2 Steps Used in Recognition of Alphabets

4.1 Image Acquisition

A camera can be used to capture the person expressing the message in sign language. Manual acquisition is possible. A camera sensor is used to capture the signer's face or gesture.

4.2 Pre-processing and segmentation

Images taken with a camera have different resolutions and sizes since they are not taken in a controlled lighting situation and are taken with a digital camera. As a result, pre-processing of images is necessary. Noise and digitization errors produce local alterations that should not significantly impact the image scene or information. Pre-processing raw video material is critical for meeting memory needs and ambient scene conditions [17].

To solve sign complexity, a variety of elements such as lighting, backdrop, camera settings, and camera location can be employed. The first and most crucial stage of pre-processing is filtering. A median filter is used to reduce unwanted noise in the recorded image. The next significant step in the pre-processing phase is substruction of background. The background subtraction are obtained using the running Gaussian average approach [18]. This method is highly fast and uses very little memory when compared to other ways. Lighting changes, camera motion adjustments, and so on are all taken into account.

Image scaling is a technique for reducing the computational effort necessary for image processing before skin detection. This technique provides a binary image where the white hand is represented by pixels and the rest of the black. This method detects whether or not each pixel in a picture is part of human skin. For skin detection in images, there are several techniques developed [20]. Different signs done with the help of hands are recognized by segmentation process. It is the process of dividing the image into pieces separated by borders. Both dynamic and static gestures impact the segmentation process. The hand movement in the case of dynamic gesture must be detected and monitored, but in the case of static gesture, the image needs to be segmented [19].

4.3 Feature Extraction

One of the most serious issue with detection system is the issue of preprocessing. Physical activity of the hands, face or entire body is used to recognise sign language. Hand gestures have a wide range of shapes, textures, and actions. Indeed, due to self-occlusion and ambient illumination, features extracted using geometric are not always available and trustworthy [21]. As stated in [22], selecting the correct feature extraction approach is essential. The system is strengthened by the combination of the recognition algorithm and feature extraction technique. In the scope of human interaction, successive hand gesture experiment is determined by the feature chosen, because the input to a classifier is gained from this stage. [23] describes the methods of the feature extraction approach.

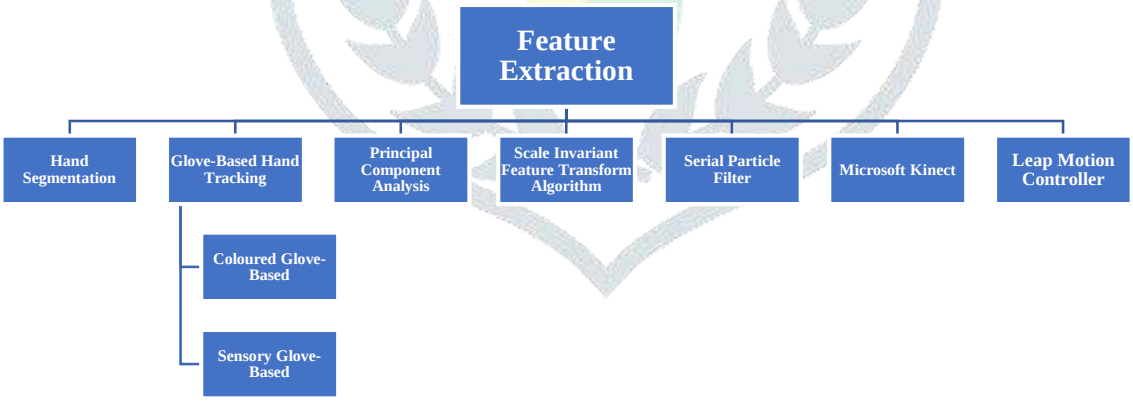

Fig. 3 Methods of Feature Extraction

4.3.1 Hand Segmentation

Segmentation is the initial stage in processing the input image for recognition. The hand area is obtained by segmenting the image that was taken by the camera. The Canny edge detector is a technique used in segmentation. As seen in this example, the hand area is detected using the canny edge detector [24]. Accuracy achieved is 92.34%. A threshold is essential when a Canny edge detector is used to locate relevant edges in an image. For example, elliptical Fourier descriptors can be employed in place of the edge detector work described in [25]. Overall accuracy is 95.1%. Skin detection is another segmentation approach. This method is used to locate the skin region in an image. This method, however, cannot calculate the hand area directly. When the facial region is included in the image, which will be detected, it may fix the problem. Skin detection must be combined with hand movement tracking in order to compute the border of hands and the centroid point of hand is calculated [26]. Another technique, as described in [27], is to separate the head and face regions by assuming that the head is typically fixed and larger than the face. To remove the face region, the face detection technique [28] is employed. The accuracy is 87.33%.

Table 2. Hand Segmentation Techniques with Accuracy

Author Name	Method	Accuracy
M. V. D. Prasad	Canny Edge Detector	92.34%
E. A. Kalsh	Canny Edge Detector	100%
P. V. V. Kishore	Elliptical Fourier Descriptors	95.1%
P. C. Pankajakshan	Skin Detection, Canny Edge Detector	-
S. C. W. Ong	Skin Detection	-
K. M. Lim	Skin Detection	87.33%

4.3.2 Glove-Based Hand Tracking

The use of gloves provided the optimum solution for simplifying the tracking of the hand area. The algorithm is aided in identifying the hands region by the usage of gloves. Gloves are frequently made of plain fabric [29]. The glove-based strategy evolves through time, moving from a coloured glove to a sensory glove approach.

A) Coloured Glove-Based

Colored gloves are one of the most traditional method for tracking hands in sign language. To identify hand movement, a hand tracking system with coloured glove is proposed [30]. It provides accuracy of 95 %. The right hand wears a bright yellow glove while the left hand wears a red and orange glove in this approach. The system looks for the proper pixel colour of the glove to determine the hand placement. After the pixel has been detected, this approach expands the testing by utilising the hand's 8 nearest neighbours to obtain the relevant part of the hand. Feature extraction is used to deliver the result as multidimensional Gaussian densities.

B) Sensory Glove-Based

Sensory gloves are equipped with seven sensors [29]. The accuracy is 88 percent. Each finger is equipped with five sensors, one sensor is used for determining hand tilt and remaining is for sensing the number of rotations. The most current innovation in sensory glove has 10 flex sensors as well as an accelerometer as a motion sensor with a 90.3 % [31]. A sensory glove might revolutionise sign recognition. It walks you through the full accelerometer pre-processing procedure. The accelerometer provides the value of hand prediction in feature vector as an input to the classification process. The proposed approach for sign language recognition when using a website employs a sensory glove as an input instruction [32]. Another study that used sensory gloves proposed Thai sign language recognition with data glove and a motion detector [33]. The accuracy is of about 94.44%. This system has been upgraded by the inclusion of 14 sensors, 10 of which are located on the fingers and 4 of which are located between the fingers.

Table 3. Glove Based Techniques with Accuracy

Author Name	Method	Accuracy
T. E. Starner	Coloured Gloves	95%
S. A. Mehdi	Sensory Gloves	88%
L. T. Phi	Sensory Gloves	90.3%
A. Ranjini S. S.	Sensory Gloves	-
S. Saengsri	Sensory Gloves	94.44%

4.3.3 Principal Component Analysis (PCA)

Principal component analysis is the well-known and widely used methodologies in scientific study. Matrixes are used in PCA's mathematical methods. PCA is being used to decrease the dimensions of a matrix, keeping just the information that is required in the matrix.

The framework [34] is composed of PCA and glove-based information collected with an Attitude Heading Reference System (AHRS) sensor. This sensor, on the other hand, does not produce a greater result. The output of that sensor is not clear and contains a lot of noise. As a consequence, the low pass filter is employed after comparing the results of the Kalman filter. This framework utilises the k-means algorithm, PCA, and an ABC-based HMM to get a acceptable result. This result had an accuracy of approximately 91.3%.

The system proposed a modified feature extraction method that combined PCA and kurtosis, as well as motion chain code (MCC) [35]. The achieved accuracy is 89.09%. Kurtosis is utilised to fuse the PCA, whereas MCC is employed to depict hand movement. Another technique [24] provided recognition of sign language by combining PCA with Elliptical Fourier Descriptors. To optimise and retain shape data, Elliptical Fourier Descriptors were used. Using Artificial Neural Network, the accuracy provided by this experiment is 92.34% in a recognition. The Otsu technique [36] was utilised for image pre-processing for segmentation in PCA. The Otsu technique output, which is in the format of a feature vector, is then compressed using PCA. For recognition, the centroid

of the hand and skin detection are used. Following table covers various proposed efforts on PCA-based algorithms, as well as the stated accuracies.

Table 4. PCA Based Techniques with Accuracy

Author	Method	Accuracy
T. H. S. Li	k-Means Algorithm, PCA	91.3%
M. M. Zaki	Kurtosis, Motion Chain Code, PCA	89.09%
M. V. D. Prasad	Elliptical Fourier Descriptors, PCA	92.34%
S. N. Sawant	Otsu Algorithm, PCA	-

4.3.4 Scale Invariant Feature Transform (SIFT) Algorithm

The system [37] proposed a SIFT-based modification feature extraction technique. The SIFT method is used by another framework [38]. This method is interesting because it is resistant to rotations, translations, and scale changes [39].

4.3.5 Serial Particle Filter

Many recognition systems employ a serial particle filter. The particle filter may track a single item, two objects, or many items [28]. The particle filter is used in sign language to track both the hands. For hand representation and recognition, a that filter is coupled with a feature covariance matrix is utilised in this study. This technique achieves an accuracy of about 87.33 % while also improving noise reduction and foreground detection. In the future, the computation time can be reduced by researching a technique for developing a virtually real-time system.

4.3.6 Microsoft Kinect

Any gesture including hand gesture is recognized with the help of Microsoft Kinect device It shapes each movement that has been recorded as a characteristic. The Microsoft Kinect is well-suited for sign language identification due to its 3D sensor camera. The camera arrangement is effective in that hands, as well as the head and elbows, are removed, which may be beneficial in differentiating the hand area and associated symptoms. Another advantage of utilising Microsoft Kinect is that it is unaffected by illumination. Even in the places with less light, camera will still detect well [40].

In the sign language recognition field, the pre-processing component has been reduced with the help of Kinect. To depict hand figures, the system offered a combination of Kinect SDK and Histograms of Oriented Gradients (HOG) [41]. The depth and coloured image were obtained using the Microsoft Kinect SDK. The shape information was normalized by HOG and feature extraction by using hand trajectory. The accuracy provided by this approach of about 96.75%.

The spatial domain method was employed as a feature extraction method in this paper to recognise the fingertips. To achieve the 97.5 % accuracy, SVM classifier was utilised [42]. The identification accuracy may be found in [43] is also of 96.67 %. It combined skeleton and depth image characteristics. The area, centroids, major and minor axes, orientations, and normalised feature vector were all considered in the feature extraction process. It was 99.84 % accurate when local and global characteristics were combined [44]. Data from Kinect was used to manage dynamic hand motions, including RGB, depth, and skeletal information. Lighting normalisation was offered as a second technique [45]. The SVM classifier produced the maximum accuracy of 99.9% [46].

Table 5. Microsoft Kinect Based Techniques with Accuracy

Author Name	Method	Accuracy
Y. Jiang	Kinect, HOG	96.75%
J. L. Raheja	Kinect	97.5%
E. Rakun	Kinect	96.67%
E. Escobedo	Kinect, Local Features, Global Features	99.84%
S. B. Carneiro	Kinect	89%
C. Keskin	Kinect	99.9%

4.3.7 Leap Motion Controller (LMC)

The active sensor was used to update the pre-processing algorithm [47]. Using a Leap Motion Controller, this system presented a feature extraction method (LMC). Using 200 frames per second, the LMC can identify hand motions. Following this, the hand movement signals are translated into computer instructions [48]. Of the 23 characteristics returned by this gadget, only 12 are selected for inclusion in the article. 15 frames per second is another way for transmitting data [49]. Various aspects of LMC's performance are superior to Microsoft Kinect according to [50]. The detected data can be directly map to the hand fingertip and

movement by using LMC API. LMC, on the other hand, is far from flawless, but it is still growing. When the hands are flipped over, it has some issues implementing the API.

4.4 Classification

The input image is classified by classifier into its related classes. During training phase, the feature vector is used to train classifier. And then trained identified class corresponding to signs and give output in text or audio format. The classifier performance measured in term of accuracy or recognition rate. Regularly used classifiers include Hidden Markov Model, Artificial Neural Network, Convolutional Neural Network, and Support Vector Machine.

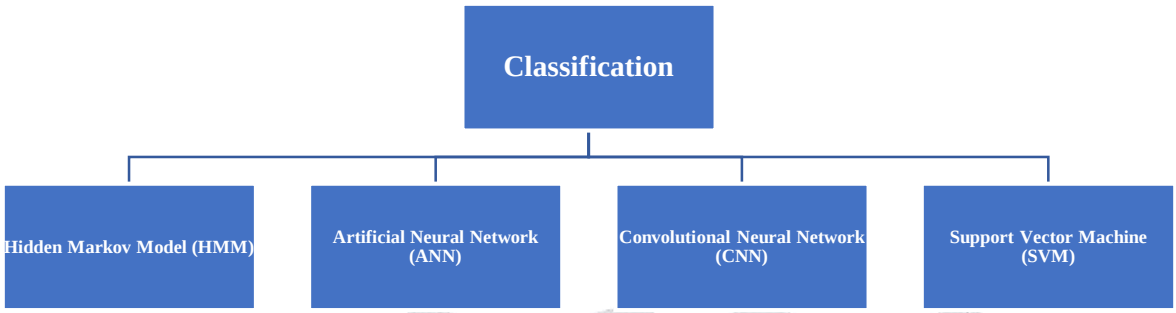

Fig.4 Classification Methods

4.4.1 Hidden Markov Model (HMM)

Multiple earlier studies use the Hidden Markov Model as a foundation for their study into building a recognition system. This is a statistical model in which a set of variables is hidden. These parameters are derived from linked observation parameters [51]. According to [52], The HMM model is a finite model that displays a probability distribution across an infinite sequence. The Hidden Markov Model is commonly used in voice recognition systems. In a glove-based SLR system, HMM is used [53]. Sign Language is made up of continuous movements that form a word or sentence. As a result, while dealing with sequential data, the HMM is used.

The HMM is multidimensional and can recognise American Sign Language with 96.7% accuracy [54]. At the same time, the input devices generated data in the form of 21 digital data, which were subsequently segmented into gesture. Following that, the data is categorised, and the result is recognised. Another study used HMM with 98.2% accuracy to recognise Chinese Sign Language (CSL). In a previous research [53], HMM was used to recognise data from Data Glove. With an accuracy of 80.4 percent, this system recognised 250 Taiwanese Sign Language phrases, which included 51 basic positions, 6 directions, and 8 key movements. Four characteristics were used to analyse the input: location, posture, motion, and orientation.

The previous research has been improved by using several Hidden Markov Model. Coupled-HMM [55], hybrid CNN-HMM [56], and Kinect-Based Hidden Markov Model [57] are used in the research for sign language recognition system. The Couple-HMM was utilised in a SLR system and achieved 90.80% accuracy [55]. In this study, the Kinect was utilised to construct a 3D representation of the collected actions. Another method [57] utilised Kinect for an HMM-based Taiwanese Language Recognition system and achieved an accuracy of 85.14%.

Table 6. HMM Used in Techniques with Accuracy

Author Name	Method	Accuracy
Wang, H.	HMM	96.7%
Liang, R. H.	HMM	80.4%
Kumar, P.	Couple-HMM	90.80%
Lee, G. C.	Kinect-based HMM	85.14%

4.4.2 Artificial Neural Network (ANN)

An artificial neural network is a component of artificial intelligence that is intended to imitate the operation of the human brain, with processing nodes known as "neurons". Every neuron contains some information; it takes input from other neurons and sends output to others [58]. With 91.1% accuracy, the ANN forward-backward method was used to recognise all alphabets and 10 numbers of ISL [1]. Backpropagation Neural Network is used in which the learning process is depends on the Deepest-Descent technique [59]. Another technique was utilised to develop a Indian Sign Language Recognition system, which obtained 92.34 % accuracy [60]. The researcher generated the data, which included 80 video sequences with 59 signs of alphabet, numerals, and 23 sentences in ISL.

Table 7. ANN Used in Techniques with Accuracy

Author Name	Method	Accuracy
Adithya, V.	ANN	96.7%
Prasad	ANN	92.30%

4.4.3 Convolutional Neural Network (CNN)

In Convolutional Neural Network, the convolutional layer extracts features [61]. CNN has been employed in studies involving computer vision. In several of these experiments, CNN was used for Sign Language Recognition. Three-dimensional CNN has been used to recognise Sign Language in written or spoken material [62]. The system's accuracy was 94.2%. The data set consists of 25 sign language words that were utilised in daily interactions. CNN was employed for an automated recognition system for Italian language that used Microsoft Kinect and GPU acceleration [63]. With a 91.7% identification rate, this system could distinguish 20 Italian gestures. A Hybrid CNN-HMM was used to recognise continuous sign language [56]. The Hidden Markov Model can be coupled with CNN to improve discriminative power. The CNN-based system obtained a recognition rate of 94.2%; however, only limited amounts of data were employed [62]. LipNet was was built with 3D Convolutional Neural Network. It is capable of lipreading using a visual method. Previously, LipNet was able to recognise end-to-end sentence-level lipreading with 95.5% accuracy. LipNet's architecture incorporates 3D CNN, Bidirectional RNNs, and SOFTMAX. The GRID corpus was used in the trials, with 3971 videos and 28775 films used for testing and training respectively [64].

Table 8. CNN Used in Techniques with Accuracy

Author Name	Method	Accuracy
Huang, J.	CNN	94.2%.
Pigou, L.	CNN, Kinect and GPU acceleration	91.7%
Koller	CNN-HMM	94.2%
Assael	CNN, LipNet	95.5%

4.4.4 Support Vector Machine (SVM)

Support Vector Machine is a pattern recognition approach that uses supervised learning. The SVM is also called as non-probabilistic binary linear classifier because it takes input data and predict that, which two possible classes to generate output. The technology [46] proposes recognising American Sign Language hand configurations of the ten digits using Kinect videos. This method creates a 3D structure of hand, which is then merged with 21 segmentation hand components to form the feature map. The experiment's SVM classifier has an accuracy of nearly 99.9%. Depends on structural and temporal data, hand signals are generated from video sequences [65]. 30 Chinese alphabet letters, with 195 photos representing each letter, for a total of 5850 images in the collection. The SVM classifier used in this system with recognition rate of 95.55%.

Conclusion

Several approaches and gesture recognition systems have been observed in the pursuit of communication channel, between the impaired hearing peoples. Excluding Indian Sign Language other language works are also discussed with the classification algorithms, challenges and the future scope were also detected. Most of the work has yet been done on the static gestures in Indian sign language. The researchers should concentrate on developing a suitable segmentation scheme capable of identifying the hand and facial area from videos and images with any environment. It has been discovered that the majority of the work must be done utilising vision-based approaches in real-time systems for sign language recognition so that it is accessible to all and provides an easy-to-use environment.

References

[1] Adithya, V., Vinod, P. and Gopalakrishnan, U. 2013. Artificial neural network based method for Indian sign language recognition. In 2013 IEEE Conference on Information & Communication Technologies, 1080–1085. Thuckalay, Tamil Nadu, India.

[2] Nanaware, T. and Sahasrabudhe, S. 2018. Fingerspelling-Indian Sign Language Trainning Tool. IEEE 18th International Conference on Advanced Learning Technologies.

[3] Patel, U and Ambedkar A. 2017. Moment Based Sign Language Recognition for Indian Language. International Conference on Computing, Communication, Control and Automation (ICCUBEA).

[4] Sajanraj T. and Beena M. 2018. Indian Sign Language Numeral Recognition Using Region of Interest Convolutional Neural Network. 2nd International conference on Inventive Communication and computational technologies.

[5] Gupta B, Shukla P and Mittal A. 2016. K-Nearest Correlated Neighbour Classification for Indian Sign Language Gesture Recognition using Feature Fusion. International Conference on Computer Communication and Informatics.

[6] Singh A, Arora S. Pushkar Shukla, Ankush Mittal, (2015) "Indian Sign Language Gesture Classification as Single Handed or Double handed" Third International Conference on Image Information Processing.

[7] S.Reshna and M.Jayaraju, (2017) " Spotting and Recognition of Hand Gesture for Indian Sign Language Recognition System with skin segmentation and SVM" International Conference on Wireless Communication, Signal Processing and Networking (WiSPNET).
[8] Sarfaraz Masood, Adhyan Srivastava, Harish Chandra Thuwal and Musheer Ahmad, (2018) "Real-Time Sign Language Gesture (Word) Recognition from Video Sequences Using CNN and RNN" Springer Nature Singapore Pte Ltd.
[9] Pradeep Kumar, Partha Pratim Roy , Debi Prosad Dogra, (2018) " Independent Bayesian classifier combination based sign language recognition using facial expression" Information Science Volume 248, 30-48.
[10] Purva C. Badhe and Vaishali Kulkarni, (2015) "Indian sign language translator using gesture recognition algorithm" IEEE International Conference on Computer Graphics, Vision and information security.
[11] Prerna Gupta, Garim Joshi, Dr. Maitreyee Dutta, (2015) "Comparative Analysis Of Movement and Tracking Techniques For Indian Sign Language Recognition" Fifth International Conference on Advanced Computing and Communication Technologies.
[12] Hardik Rewari, Vishal Dixit, Dhroov Batra, Hema N, (2018) "Automated Sign Language Interpreter" Proceedings of Eleventh International Conference on Contemporary Computing(IC3), 2-4 August. 2018, Noida, India.
[13] Kumud Tripathi, Neha Baranwal, G.C. Nandi, (2015) " Continuous Dynamic Indian Sign Language Gesture Recognition With Invariant Backgrounds" International Conference on Advances in Computing, Communication and Informatics (ICACCI).
[14] Shwetha S kulloli, R H goudar, (2017) " A Effective Communication Solution for hearing Impaired Persons: A novel Approach using gesture and sentence Formation " International Conference On Smart Technology for Smart Nation.
[15] Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar, (2018) "Real-time Indian Sign Language(ISL) Recognition" 9th ICCCNT, 10-12, IISC, Bengaluru,India.
[16] G.Ananth Rao, P.V.V. Kishore, (2016)"Sign Lanuage Recognition System Simulated for Video Captured with Smart Phone Front Camera", International Journal Of Electrical and Computer Engineering, 6(5), 2176-2187.
[17] Brian L. Pulito, Raju Damarla, Sunil Nariani, (1990) " 2-D Shift Invariant image Classification Neural Network, which overcomes Stability, Plasticity Dilemma", Vol 2, International Joint Conference on Neural Network, San Deigo.
[18] Jong Bae Kim,Hye Sun Park,Min Ho Park,Massimo Piccardi, (2004) 'Background subtraction techniques: a review',Systems, Man and Cybernetics,vol.4,IEEE InternationalConference,3099-3104,.
[19] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012) "comparative study of skin color based segmentation techniques", Aligarh Muslim University A.M.U., Aligarh, India.
[20] Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva, (2002)" A Survey on Pixel-Based Skin Color Detection Techniques", Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University.
[21] J. P. Wachs, H. Stern, and Y. Edan, (2005) "Cluster labeling and parameter estimation for the automated setup of a hand-gesture recognition system," IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 35(6), 932–944.
[22] N. K. Gilorkar and M. M. Ingle, (2014) "A review on feature extraction for Indian and American sign language," International Journal of Computer Science and Information Technologies, 5(1), 314–318.
[23] H. Bhavsar, (2017) "Review on feature extraction methods of image based sign language recognition system," Indian Journal of Computer Science and Engineering, 8(3), 249–259.
[24] M. V. D. Prasad, V. Kishore, and A. Kumar, (2016) "Indian sign language recognition system using new fusion based edge operator," Journal of Theoretical and Applied Information Technology, 88(3), 574–584.
[25] P. V. V. Kishore, M. V. D. Prasad, C. R. Prasad, and R. Rahul, (2015) "4-camera model for sign language recognition using elliptical Fourier descriptors and ANN," International Conference on Signal Processing and Communication Engineering Systems (SPACES), IEEE, pp. 34–38.
[26] P. C. Pankajakshan and B. Thilagavathi, (2015) "Sign language recognition system," Innovations in Information, Embedded and Communication Systems (ICIIECS), IEEE, pp. 1–4.
[27] S. C. W. Ong and S. Ranganath, (2005) "Automatic sign language analysis: A survey and the future beyond lexical meaning," IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(6), 873–891.
[28] K. M. Lim, A. W. C. Tan, S. C. Tan, (2016) "A feature covariance matrix with serial particle filter for isolated sign language recognition," Expert Systems with Applications, Elsevier Ltd, 54, 208–218.
[29] S. A. Mehdi and Y. N. Khan, (2002) "Sign language recognition using sensor gloves," Proceedings of the 9th International Conference on Neural Information Processing (ICONIP), IEEE, vol.5, 2204–2206.
[30] T. E. Starner, (1995) "Visual recognition of American sign language using hidden Markov models," Massachusetts Inst of Tech Cambridge Dept Of Brain And Cognitive Science.
[31] L. T. Phi, H. D. Nguyen, T. Q. Bui, T. T. Vu, (2015) "A glove-based gesture recognition system for Vietnamese sign language," Proceedings of the 15th International Conference on Control, Automation and Systems, 1555–1559.
[32] A. Ranjini S. S. and M. Chaitra, (2014) "Sign language recognition system," International Journal on Recent and Innovation Trends in Computing and Communication, 2(4), 947–953.
[33] S. Saengsri, V. Niennattrakul, and C. A. Ratanamahatana, (2012) "TFRS: Thai finger-spelling sign language recognition system," 2nd International Conference on Digital Information and Communication Technology and its Applications (DICTAP), 457–462.
[34] T. H. S. Li, M. C. Kao, and P. H. Kuo, (2016) "Recognition system for home-service-related sign language using entropy-based k-means algorithm and ABC-based HMM," IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46(1), 150–162.
[35] M. M. Zaki and S. I. Shaheen, (2011) "Sign language recognition using a combination of new vision based features," Pattern Recognition Letters, Elsevier B.V., 32(4), 572–577.
[36] S. N. Sawant, (2014) "Sign language recognition system to aid deaf-dumb people using PCA," International Journal of Computer Science and Engineering Technology, 5(5), 570–574.

[37] Thorat, V. Satpute, A. Nehe, T. Atre, and Y. R. Ngargoje, (2014) "Indian sign language recognition system for deaf people," International Journal of Advanced Research in Computer and Communication Engineering, 3(3), 5319–5321.
[38] S. Goyal, I. Sharma, and S. Sharma, (2013) "Sign language recognition system for deaf and dumb people," International Journal of Engineering Research and Technology, 2(4), 382–387.
[39] P. Gurjal and K. Kunnur, (2012) "Real Time Hand Gesture Recognition Using SIFT," International Journal of Electronics and Electrical Engineering, 2(3), 19–33.
[40] M. A. Majid and J. M. Zain, (2013) "A review on the development of indonesian sign language recognition system," Journal of Computer Science, 9(11), 1496–1505.
[41] Y. Jiang, J. Tao, W. Ye, W. Wang, and Z. Ye, (2014) "An isolated sign language recognition system using RGB-D sensor with sparse coding," 17th International Conference on Computational Science and Engineering, IEEE, 21–26.
[42] J. L. Raheja, A. Mishra, and A Chaudhary, (2016) "Indian sign language recognition using SVM," Pattern Recognition and Image Analysis, 26(2), 434–441.
[43] E. Rakun, M. Andriani, I. W. Wiprayoga, K. Danniswara, and A. Tjandra, (2013) "Combining depth image and skeleton data from kinect for recognizing words in the sign system for Indonesian language (SIBI [sistem isyarat bahasa Indonesia])," International Conference on Advanced Computer Science and Information Systems (ICACSIS), 387–392.
[44] E. Escobedo and G. Camara, (2016) "A new approach for dynamic gesture recognition using skeleton trajectory representation and histograms of cumulative magnitudes," SIBGRAPI Conference on Graphics, Patterns and Images, 209–216.
[45] S. B. Carneiro, E. D. D. M. Santos, M. D. A. Talles, J. O. Ferreira, S. G S. Alcala, and A. F. Da Rocha, (2016) "Static gestures recognition for Brazilian sign language with Kinect sensor," SENSORS, IEEE.
[46] C. Keskin, F. Kırac, Y. E. Kara, and L. Akarun, (2011) "Real time hand pose estimation using depth sensors," International Conference on Computer Vision Workshops, IEEE, 1228–1234.
[47] M. Mohandes, S. Aliyu, and M. Deriche, (2014) "Arabic sign language recognition using the leap motion controller," 23rd International Symposium on Industrial Electronics (ISIE), IEEE, 960–965.
[48] H. Bhavsar, (2017) "Review on classification methods used in image based sign language recognition system," International Journal on Recent and Innovation Trends in Computing and Communication, 5(5), 949–959.
[49] M. Koul, P. Patil, V. Nandurkar, and S. Patil, (2016) "Sign language recognition using leap motion sensor," International Research Journal of Engineering and Technology (IRJET), 3(11), 322–325.
[50] L. E. Potter, J. Araullo, and L. Carter, (2013) "The leap motion controller: A view on sign language," Proceedings of the 25th Australian Computer-Human Interaction, Conference on Augmentation, Application, Innovation, Collaboration – OzCHI, ACM, 175–178.
[51] Zabulis, X., Baltzakis, H., & Argyros, A. (2009). Vision-based hand gesture recognition for human-computer interaction. In The universal access handbook, 1–30. Boca Raton: CRC Press.
[52] Eddy, S. R. (1996). Hidden markov models. Current Opinion in Structural Biology, 6(3), 361–365.
[53] Liang, R. H., & Ouhyoung, M. (1998). A real-time continuous gesture recognition system for sign language. In Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 558–567). Nara, Japan.
[54] Wang, H., Leu, M. C., & Oz, C. (2006). American Sign Language Recognition Using Multi-dimensional Hidden Markov Models. Journal of Information Science and Engineering, 22(5), 1109–1123.
[55] Kumar, P., Gauba, H., Roy, P. P., & Dogra, D. P. (2017). Coupled hmm-based multi-sensor data fusion for sign language recognition. Pattern Recognition Letters, 86, 1–8.
[56] Koller, O., Zargaran, O., Ney, H., & Bowden, R. (2016). Deep Sign: Hybrid CNN-HMM for Continuous Sign Language Recognition. In Proceedings of the British Machine Vision Conference 2016. New York, UK.
[57] Lee, G. C., Yeh, F.-H., & Hsiao, Y.-H. (2016). Kinect-based taiwanese sign-language recognition system. Multimedia Tools and Applications, 75(1), 261–279.
[58] Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. International Journal of Engineering and Innovative Technology, 2(1), 189–194.
[59] Buscema, M. (1998). Back propagation neural networks. Substance Use & Misuse, 33(2), 233–270.
[60] Prasad, D., Kishore, V., & Kumar, A. (2016). Indian sign language recognition system using new fusion based edge operator. Journal of Theoretical and Applied Information Technology, 88(3), 574–584.
[61] Pham, N. Q., Kruszewski, G., & Boleda, G. (2016). Convolutional Neural Network Language Models. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing 1153–1162. Austin, Texas.
[62] Huang, J., Zhou, W., Li, H., & Li, W. (2015). Sign language recognition using 3D convolutional neural networks. In Multimedia and Expo (ICME), 2015 IEEE International Conference on 1–6. Turin, Italy.
[63] Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2014). Sign language recognition using convolutional neural networks. In Workshop at the European Conference on Computer Vision, 572–578. Cham : Springer.
[64] Assael, Y. M., Shillingford, B., Whiteson, S., & de Freitas, N. (2016). LipNet: Sentence-level Lipreading. Computing Research Repository (CoRR), 1611.01599, 1-13. Retrieved July 4, 2017 from http://arxiv. org/abs/1611.01599.
[65] Y. Quan, "Chinese sign language recognition based on video sequence appearance modeling," in Proc. 5th IEEE Conf. Industrial Electronics and Applications (ICIEA), 2010, 1537–1542.