A proposed framework for Sign Language Recognition

Abhisar singh Raghuvanshi , Anshika Sharma & Khushi Gupta

Department of Computer Science & Engineering(Artifical Intelligence ) ,UIET Kanpur

**ABSTRACT**: Sign language recognition is a topic of current research in Computer Science and Engineering field. This application will be a boost to the deaf and hard hearing people. They are not able to use the computers and other hand held devices as it is very difficult for them to interact with such devices. So, in this area a lot of research is going on to help them. In India, either no standard database is available to carry research in this area and or no systems are available for them. In the proposed research we will try to develop a system for hard hearing and physically challenged persons. Here we present an overview of current research dimension in national and international scenario.

**Keywords**: Sign Language, Indian Sign Language (ISL), Pattern Recognition

**INTRODUCTION**

Sign language (SL) [1] is a visual-gestural language used by deaf and hard-hearing people for communication purposes. They use three dimensional spaces and the hand movements (and other parts of the body) to convey meanings. It has its own vocabulary and syntax entirely different from spoken languages. Spoken languages use the oratory faculties to produce sounds mapped against specific words and grammatical combinations. The oratory elements are then received by the auditory faculties and processed accordingly. Sign language rather uses the visual faculties. Spoken language makes use of rules to produce comprehensive messages whereas sign language is also governed by a complex grammar. The Sign Language development is different for each country or sub-continent. The following table presents the development of sign languages of influencing countries/subcontinent. The table below highlights the similarities and differences in their sign languages..

…………………………………………………………………………………………………………………………………………………………….......

**Sign Language**

Professionals in India believe in an acute shortage of special schools for deaf people. A very few schools use sign language as a medium of instruction. There is also a lack of proper and effective audio visual support in oral education in these schools. This results in inadequate communication and language skills in the majority of deaf children, impacting on poor literacy skills in the deaf community. The reality is that deaf schools mainly do not use ISL and nearly 5% of deaf people [12, 13, 16] attend deaf schools. The use of ISL is restricted only to vocational programs and short term courses. ISL was partly influenced by BSL in the finger spelling system and some other signs, but most are unrelated to European sign system. There was no formal ISL until 1978. Banerjee [14] compared signs used in some schools for the deaf situated in West Bengal and part of Assam. His conclusion was that gestures used in each school were not the same. He believed that signing started in India in the 18th century but its use was strongly discouraged. Madan Vasishta [17] sent a questionnaire to the heads of more than hundred schools for the deaf in India in 1975. Almost all the respondents agreed that there was no ISL. But they also acknowledged that deaf children used some kind of gestures. A similar survey was conducted 20 years later, using questionnaires sent out to schools for the deaf. Some of the responses show the same misconceptions about sign language that signing is “based on spoken language”, or “based on English”, or “difficult to provide a sign for every spoken word”. Some statements showed that a more positive attitude towards manual communication, and here respondents talked about sign language, rather than gestures. Increasing awareness about the nature of sign languages is evidenced later on.

**PROBLEM DEFINITION**

Observing the advantages of works on Sign Language recognition of different countries in aiding the deaf people for communication in public places and accessing/communicating with latest gadgets like Telephone, Computers, etc, linguistic studies on Indian Sign Language started in 1978 in India. These works resulted in ISL and discovered that it is a language on its own with specific syntax, grammar, phonology and morphology. While significant progress has already been made in computer recognition of sign languages of other countries but a very limited work has been done in ISL computerization [19]. The sign language recognition produces sentences in machine readable form so it can be translated in a desired form using Machine Translation [MT] system and vice-versa. Most of the MT systems use one of the following three basic approaches. Direct translation• Transfer-based Architecture• Interlingua-based Architecture• In the proposed research work our sign recognition system will recognize sign language and translate into a form acceptable by a system based on interlingua-based architecture. In turn this architecture would represent in an internationally acceptable form so that systems for other languages can be interfaced.

………………………………………………………………………………………………………………………………………………………………...

**DETAILED METHODOLOGY**

Automatic recognition of a sign language will track three target objects, the face and the two hands, and the extraction of features which will be classified. Tracking is a difficult task since the face and hands are of the same color and frequently overlap from a viewing point or touch. Thus the identification and the segmentation of occluded objects are necessary for the purpose of feature extraction. Features specify signs using the global representation that deals with motion trajectories and coarse shapes of the hands, or the local representation that deals with the characteristics of the fine hand shapes. These features will then to be classified as signs in the recognition process. The signs in the vocabulary will be modeled through training within the selected feature space, and used for classification.

Methodologies to be used in the research work can be divided into the following phases:

**1 Data collection**

The Indian sign data required in this research are to be collected from different deaf schools situated in India. As described by Zeshan et al [12] standardized Indian sign gestures in the form of image and video can be obtained from Ali Yavar Jung National Institute for Hearing Handicapped (AYJNIHH).

**2 Critical review of data**

The collected data are to be reviewed critically to form a standard database. The review process is progress.

**3 Pattern recognition**

Pattern recognition stems from the need for automated machine recognition of objects, signals or images, or the need for automated decision-making based on a given set of parameters. Despite over half a century of productive research, pattern recognition continues to be an active area of research because of many unsolved fundamental theoretical problems as well as a rapidly increasing number of applications that can benefit from pattern recognition. The following table represents various methods used in pattern recognition that we will use in our research work.

**4 Feature extraction**

Feature extraction is an essential pre-processing step to pattern recognition and machine learning problems. It is often decomposed into feature construction and feature selection. Most likely methods may be used in the research are:

o Statistical Features

o Structural Features

o Hybrid Features

**5 Classification**

In this module some of the classifications of systems will be briefly executed and the most important properties of these systems are to be explored. Understanding these basic differences between systems, and their properties, will be a fundamental concept used in gesture recognition.

**6 Testing of data**

The standard data are to be used in sign language recognition. The result of various recognition techniques are then to be compared to find out the better among them.

**7 Industry grade software to be developed**

A software system for recognizing Indian Sign Language is to be developed so that the deaf community in India can use it for communicating with different gadgets. For the purpose the software as well as interfaces to different gadgets and other sign languages are to be developed. We will also try to integrate it with mobile phones and other hand held devices where the system will be helpful at any public place.

………………………………………………………………………………………………………………………………………………………………….

**EXPECTED OUTCOMES**

The following are expected outcomes of the proposed research:

• Standard database for ISL gestures

• A computer based system for ISL recognition and translation

• Patent on our database

• Patent on our system

**RESULTS AND DISCUSSION**

Here, we give specifics on the datasets that were captured using the suggested SLR methodology. The continual identification of sign language has then been given. The outcomes of the word recognition of solitary signs have finally been presented. Four people have signed up to participate in the collection of sign language data. Two of the participants were young people who attended a hearingimpaired school in Gowa, Indonesia. They are from the Indonesian city of Makassar's Community Deaf. 20 solitary signs in words make up the dataset. Each signatory has said each sign word at least 30 times. As a result, 2400 sign words (20\*30\*4) in all are noted. In Table I, you'll find a comprehensive description of all the sign words, with the '$' sign indicating transitional movements, representing the switch between two continuous letters within a sign sentence. Twenty examples of dynamic gesture visualization are presented to justify this model. They were selected based on the words most frequently used by deaf people, namely: "Love", "Everything", "Ball", "Rice", "There", "Remember", "They"," Market"," You"," Hear"," Play"," Go"," See"," Laugh"," We"," Honestly"," Cry"," Sleep"," Cook" and "$". Table I shows visualizations of the selected words.

TABLE I

WORDS SIGN LANGUAGE IN THE DATASET

Love Everything Ball There Remember They Market You Hear Go See Laugh We Honestly Cry Sleep Cook Play

In Table 1, twenty words in the dataset represent the categories of nouns, verbs, adjectives, and adverbs. In addition, the words selected in the data set are the most frequently used words by deaf people in their daily activities. The purpose of the word selection is to be able to make a sentence that is intact for the deaf at the time of the test. In Figure, an illustration of a dynamic signal on a sentence sample is provided.

This research makes a valuable contribution to the Human Action Recognition (HAR) field, showcasing the potential of utilizing MediaPipe and modified LSTM for feature extraction and model training. The obtained accuracy levels, particularly in word recognition, underscore the success of this method in addressing HAR challenges.

**Mediapipe**

Sign language recognition has the potential to improve the situation of a large number of disabled people while dealing with normal human beings but the use of sign language for communication is limited. As a result, there is a need to create a more convenient approach for persons with hearing impairments to learn and work in order to improve their lives.

Gesture recognition has been studied extensively utilizing traditional techniques such as body component tracking, different color glove-based tracking, Kinect depth sensor tracking, and skeleton tracking. Multiple methods have been used to solve this problem like modified CNN, image segmentation, SVM and deep learning. Many machine learning algorithms have been developed for hand gesture recognition so as to create AI-based applications. Out of them, MediaPipe can be used for hand gesture recognition.

Google supported MediaPipe framework can be used for solving several problems like face-recognition, face-map, eye, hand, poseestimator, holistic, hair, object-detection, box tracking and KIFT. With the help of the MediaPipe framework, we can develop an algorithm or model for the application, then help the application by providing results that can be cloned across different platforms. The MediaPipe framework is composed of three major components:

(1) performance evaluation,

(2) a mechanism for collecting data from the sensor

(3) an assembly of reusable parts.

A graph consisting of all the parts called the calculators is known as pipeline, wherein every calculator is inter-connected by channels through which the data flows. Developers can create their required application by removing or delineating user defined calculators anywhere in the graph. This result of calculators and channels creates a data-flow diagram.

Hand gesture recognition with the MediaPipe framework is a dependable and high-fidelity hand and fingertracking system. Mediapipe hands uses an integrated ML pipe of several models working together [18]:

(1) A palm recognizer processes the captured hand image,

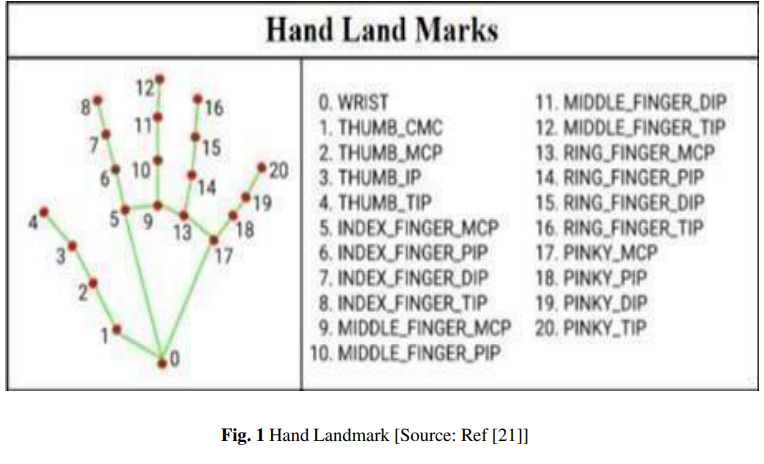
(2) A hand landmark model takes processed image as input and returns hand with 3D key points as output.

(3) A gesture recognition model which processes the 3D hand key-points and classifies them into a discrete set of gestures.

[4] The palm detection model outputs a precisely cropped picture of the palm that is then sent to the landmark model. This method does away with data augmentation, which is used in deep learning models

[5] to rotate, flip, and scale images. The technique of detecting hands is time-consuming and difficult since it involves working with different hand sizes, thresholding, and image processing. Prior to identifying hands with connected fingers, a palm detector is trained, which estimates bounding boxes around hard objects like fists and the palm. The second method is to utilize an encoder-decoder as an extractor for a larger scene context [14]. Hand Landmark model implements machine learning model to take 21 3-D key points of a hand from just a frame using regression which will directly produce the coordinate prediction. Even with faintly visible hands and self-occlusions, the model acquires a rigorously defined hand position representation

[6]. It provides better real-time performance on devices compared to other algorithms and can be scaled for multiple-hands in a single frame.



**CONCLUSION**

In this study, we have developed a novel framework for sustainable SLR employing Mediapipe and cameras. The addition of sign and sentence words has also been suggested for the modified LSTM architecture. A data set of 15 distinct sign words was utilized to train the model. Our assessment of the technique is based on 320 signed sentences written by four signatories. On signed sentences and isolated marked words, average accuracy levels of 80.0% and 85% have been noted. In the future, additional training data for greater model learning can enhance introduction performance and get a new method for the detection transition sign.

Individuals with hearing disabilities often face significant challenges in communicating with people who can hear. One of the most effective ways for them to communicate is through sign language. However, for people who do not know sign language, understanding what is being communicated can be a significant challenge. This communication gap can have a detrimental impact on the social and emotional well-being of individuals with hearing disabilities, making it difficult for them to engage fully in society.

Although there are still some research gaps that need to be addressed, such as improving the system's accuracy in recognizing signs for complex phrases and developing a portable and affordable device for practical use in daily life, the proposed Sign Language Recognition system offers a promising step towards creating a more inclusive society. With further development and refinement, this system can play a significant role in breaking down communication barriers and facilitating greater accessibility and understanding for individuals with hearing disabilities.

**ACKNOWLEDGMENT**

The authors thank the Ministry of Education, Culture, Research and Technology of the Republic of bharat, UIET KANPUR, CSJM UNIVERSITY for provision of the research facilities.

**REFERENCES**

[1] B. Sundar and T. Bagyammal, "American Sign Language Recognition for Alphabets Using MediaPipe and LSTM," Procedia Comput Sci, vol. 215, pp. 642–651, 2022, doi:10.1016/j.procs.2022.12.066.

[2] Y. SHI, Y. LI, X. FU, M. I. A. O. Kaibin, and M. I. A. O. Qiguang, "Review of dynamic gesture recognition," Virtual Reality and Intelligent Hardware, vol. 3, no. 3. KeAi Communications Co., pp. 183–206, Jun. 01, 2021. doi:10.1016/j.vrih.2021.05.001.

[3] D. K. Jain, A. Kumar, and S. R. Sangwan, "TANA: The amalgam neural architecture for sarcasm detection in indian indigenous language combining LSTM and SVM with word-emoji embeddings,"Pattern Recognit Lett, vol. 160, pp. 11–18, Aug. 2022, doi:10.1016/j.patrec.2022.05.026.

[4] Y. S. Tan, K. M. Lim, and C. P. Lee, "Hand gesture recognition via enhanced densely connected convolutional neural network," Expert Syst Appl, vol. 175, Aug. 2021, doi:10.1016/j.eswa.2021.114797.

[5] P. K. Athira, C. J. Sruthi, and A. Lijiya, "A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 3, pp. 771–781, Mar. 2022, doi:10.1016/j.jksuci.2019.05.002.

[6] R. O. Maimon-Mor et al., "Talking with Your (Artificial) Hands: Communicative Hand Gestures as an Implicit Measure of Embodiment," iScience, vol. 23, no. 11, Nov. 2020, doi:10.1016/j.isci.2020.101650.

[7] V. Adithya and R. Rajesh, "Hand gestures for emergency situations: A video dataset based on words from Indian sign language," Data Brief, vol. 31, Aug. 2020, doi: 10.1016/j.dib.2020.106016.

[8] A. P. G and A. P. k, "Design of an integrated learning approach to assist real-time deaf application using voice recognition system," Computers and Electrical Engineering, vol. 102, p. 108145, Sep. 2022, doi:10.1016/j.compeleceng.2022.108145.

[9] C. Hinchcliffe et al., "Language comprehension in the social brain: Electrophysiological brain signals of social presence effects during syntactic and semantic sentence processing," Cortex, vol. 130, pp. 413–425, Sep. 2020, doi:10.1016/j.cortex.2020.03.029.

[10] M. Suneetha, P. MVD, and K. PVV, "Multi-view motion modelled deep attention networks (M2DA-Net) for video based sign language recognition," J Vis Commun Image Represent, vol. 78, p. 103161, Jul. 2021, doi:10.1016/J.JVCIR.2021.103161.

[11] K. Sadeddine, Z. F. Chelali, R. Djeradi, A. Djeradi, and S. Ben Abderrahmane, "Recognition of user-dependent and independent static hand gestures: Application to sign language," J Vis Commun Image Represent, vol. 79, p. 103193, Aug. 2021, doi:10.1016/j.jvcir.2021.103193.

[12] L. R. Cerna, E. E. Cardenas, D. G. Miranda, D. Menotti, and G. Camara-Chavez, "A multimodal LIBRAS-UFOP Brazilian sign language dataset of minimal pairs using a microsoft Kinect sensor," Expert Syst Appl, vol. 167, p. 114179, Apr. 2021, doi:10.1016/j.eswa.2020.114179.

[13] R. Solgi, H. A. Loáiciga, and M. Kram, "Long short-term memory neural network (LSTM-NN) for aquifer level time series forecasting using in-situ piezometric observations," J Hydrol (Amst), vol. 601, Oct. 2021, doi:10.1016/j.jhydrol.2021.126800.

[14] L. Gao, H. Li, Z. Liu, Z. Liu, L. Wan, and W. Feng, "RNN-Transducer based Chinese Sign Language Recognition," Neurocomputing, vol. 434, pp. 45–54, Apr. 2021, doi:10.1016/j.neucom.2020.12.006.

[15] S. Subburaj and S. Murugavalli, "Survey on sign language recognition in context of vision-based and deep learning," Measurement: Sensors, vol. 23, p. 100385, Oct. 2022, doi:10.1016/j.measen.2022.100385.

[16] K. Anand, S. Urolagin, and R. K. Mishra, "How does hand gestures in videos impact social media engagement - Insights based on deep learning," International Journal of Information Management Data Insights, vol. 1, no. 2, Nov. 2021, doi:10.1016/j.jjimei.2021.100036.

[17] R. Solgi, H. A. Loáiciga, and M. Kram, "Long short-term memory neural network (LSTM-NN) for aquifer level time series forecasting using in-situ piezometric observations," J Hydrol (Amst), vol. 601, Oct. 2021, doi:10.1016/j.jhydrol.2021.126800.

[18] P. K. Athira, C. J. Sruthi, and A. Lijiya, "A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 3, pp. 771–781, Mar. 2022, doi:10.1016/j.jksuci.2019.05.002.

[19] G. A. Rao and P. V. V. Kishore, "Selfie video based continuous Indian sign language recognition system," Ain Shams Engineering Journal, vol. 9, no. 4, pp. 1929–1939, Dec. 2019, doi:10.1016/j.asej.2016.10.013.

[20] S. K. Devi and S. CN, "Intelligent Deep Learning Empowered Text Detection Model from Natural Scene Images," Int J Adv Sci Eng Inf Technol, vol. 12, no. 3, pp. 1263–1268, 2022, Accessed: Nov. 07, 2023. [Online]. doi:10.18517/ijaseit.12.3.15771.

[21] R. Sreemathy, M. P. Turuk, S. Chaudhary, K. Lavate, A. Ushire, and S. Khurana, "Continuous word level sign language recognition using an expert system based on machine learning," International Journal of Cognitive Computing in Engineering, vol. 4, pp. 170–178, Jun. 2023, doi:10.1016/J.IJCCE.2023.04.002.