IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

# INDIAN SIGN LANGUAGE INTERPRETER FOR DEAF AND MUTE PEOPLE

[1]Aditi Chavan, [2]Jayshree Ghorpade-Aher, [3]Aakriti Bhat, [4]Aniket Raj, [5]Shubham Mishra
[1]Student, [2]Professor, [3-5]Student
[1]Computer Science and Technology,
[1]Dr. Vishwanath Karad Maharashtra Institute of Technology World Peace University, Pune, India

***Abstract:*** Sign Language is the primary mode of interaction between the specially-abled deaf and mute population. As per the World Health Organization (WHO) survey, 460 million individuals including men, women and children suffer hearing disability. Out of these 460 million individuals, about 12.3 million belong to India. Most of this population relies on a third person who is capable of translating hand sign gestures to the regionally spoken language for them. But involving a translator hinders the conversation privacy. Here, the rapid development in the field of Artificial Intelligence (AI) and Machine Learning (ML) aids to convert hand sign gestures into words that can be understood by normal people. The Indian sign language interpreter for the mute and deafened people follows a Vision-based approach which uses the Machine Learning technique of 3D-CNN (3-Dimensional Convolutional Neural Network) and LSTM (Long Short-Term Memory) neural network to productively map the input-Indian sign language gesture images to their meaning. This system aims to interpret hand gestures images of Indian Sign Language for alphabets, numbers, sentences as well as emergency words (conversion of images of hand gesture to text/audio). This proposed technique with data from various sources will aid to alleviate the communication gap between the deaf and mute people and others.

***Index Terms* - 3D Convolutional Neural Network (3D CNN), Faster R-CNN (Faster Recurrent Neural Network), Indian Sign Language (ISL), Long Short-Term Memory Neural Network (LSTM), Sign Language Interpreter (SLI)**

## I. INTRODUCTION

Communication is the most important part of living. Sign language is the only medium for deaf and mute people to communicate but due to its triviality and insufficient knowledge of different Sign Languages based on regions, it becomes hard for them to communicate which is why many of them experience communication issues in day-to-day life.

**Global Statistics -** As per the World Health Organization (WHO) survey, over 460 million individuals which constitutes about 5% of the world population have hearing disabilities.
**Indian Statistics -** From these 460 million individuals about 12.3 million belong to India. Communicating with this specially-abled population is difficult because not everyone is well-versed with the sign language gestures.

**Rights of Persons with Disabilities Act 2016 (RPwD Act 2016) -** The RPwD Act 2016 resolute appointing ISL interpreters at all Government and public sector organizations in order to create an unhindered communicating environment for the disabled people.

**Existing solution -** A recognized solution for Sign Language Interpretation is a manual sign language translator (an individual who is skilled to translate hand sign gesture to words and vice-versa), but this solution hinders privacy of the conversation. For an Indian deaf and mute population of 1.8 million to 7 million, the number of certified Sign Language Interpreters in the country is only 250 and the count goes on decreasing due to reluctancy or lack of qualification to pursue the job.

Here lies the need to develop an automated sign language interpreter enabled to convert the hand sign language to alphabetical language which will help in overcoming the communication barrier between the specially-abled deaf and mute people and us. The automated interpreter can either be Vision-based or Glove/Sensor based.

The proposed system of The Indian Sign Language Interpreter for Deaf and Mute people, interprets the hand sign gesture images [19] for basic Banking Help-desk words such as open account, fill passbook, withdraw money, amount value (like Rs. 5,000), etc. Allows the user to read out the interpretation (text to speech conversion of the interpreted meaning). Additional option to assign one's own meaning to sign gesture (to avoid complex gestures. Although the main dataset of ISL won't be altered, this option will create a user's dictionary).

## II. LITERATURE SURVEY

### 2.1 Overview

The Vision-based Approach [1-17] deals with the interpretation of 2D or 3D vision. In 2D sensing the width and height of the gesture are taken into consideration, whereas in 3D sensing the depth factor is also required. 3D technique requires the use of depth sensing cameras. While forming some hand gestures it is possible that our fingers might overlap, hence the depth sensing feature of 3D sensing yields better results. The 2D/3D sensing combined with AI/ML - Neural Network algorithms help build a Sign Language Interpreter.

Another approach being the Glove/Sensor-based [19-21] approach consist of three parts namely- input, processing and output units. In the Input unit, sensors are there for measuring hand and finger movements. Most commonly used sensors include accelerometers, proximity sensors, inertial measurement unit (IMU), flexion sensors and abduction sensors. Processing unit acts as a microcontroller that utilizes the data collected from sensors for processing and recognition. The output unit is usually a GUI or speakers.

### 2.2 Related Research Work

Sruthi C. J and Lijiya A [2] described a vision-based approach for static alphabet recognition of Indian Sign Language (ISL). The model used Viola-Jones face detection technique to identify faces of the signer and thus eliminate it, followed by Skin HSV (Skin color segmentation algorithm) to extract the hand gesture region. The CNN used, helped map the input binary silhouette to its interpretation. The model successfully achieved an accuracy of 98.64% for static alphabet recognition. It couldn't perform finger-spelling dynamic alphabets and dynamic words.

Siming He [7] proposed a Vision-based system which dealt with video frames as input for Sign language recognition. The system used Faster R-CNN to detect hand gestures from video input and also gave a comparison of Faster R-CNN w.r.t. Fast R-CNN and YOLO. Dual channel 3D CNN was used to maintain the video frames as well as its sequence during feature extraction. Finally, LSTM Neural Network encoded and decoded the features extracted by 3D CNN and mapped them to their meaning. Accuracy rates reached 99.0% and the model used an RGB data stream for a limited dataset.

This paper [4] helps to get the meaning of sign language in text format by considering the dataset of Indian Sign Language and taking input in image format using a camera. Firstly, Image processing is done to extract features. HSV is the method they have selected for feature extraction. HSV stands for Hue, Saturation, and Value. These values are selected according to the background in such a way that the hand gets isolated from the background in order to get the hand gestures correctly. Then the images gathered from HSV are compared from the datasets and Deep learning is implemented to interpret the gestures. The classification process based on a 3-layer CNN network helps interpret the hand sign gesture.

The novel approach used in paper [10] consists of two modules i.e., Sign to Text (STT) and Text to sign. Firstly, in this sign language video captured is pre-processed to detect motion frames and reduce the number of frames to be processed. Then captured frames are sent to trained models to detect the sign in the image. Next to find out the sign from text or phrase it is passed as an input to the model. Here phrases are split into words, since articles and helping words do not have signs so they are ignored, then images for words are fetched from the cloud and are displayed. STT faces a lot of challenges which includes appropriate pattern matching due to presence of noise in images. This literature uses fusion of 3D depth and RGB data to capture images so as to increase the efficiency of the model.

## III. PROPOSED METHODOLOGY

The proposed method for SLI is designed on the vision-based approach, wherein the input to the system is a hand gesture image captured through a camera and the model maps the image to its interpretation using Machine Learning (ML) techniques specified in Section 3.2.

### 3.1 System Architecture

Figure 1 depicts the SLI architecture consisting of four sequential processes - hand gesture region extraction, AI/ML model of 3D CNN for extracted feature handling and then the LSTM encoding and decoding to gesture interpretation.

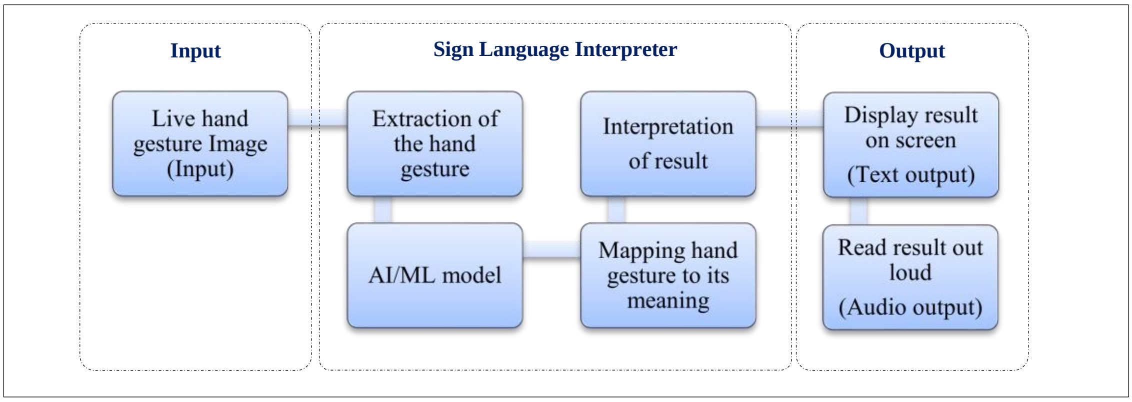


Figure 1: System Architecture of SLI

### 3.2 Technical Modules

The modules of the proposed system are elaborated as shown in Figure 2.

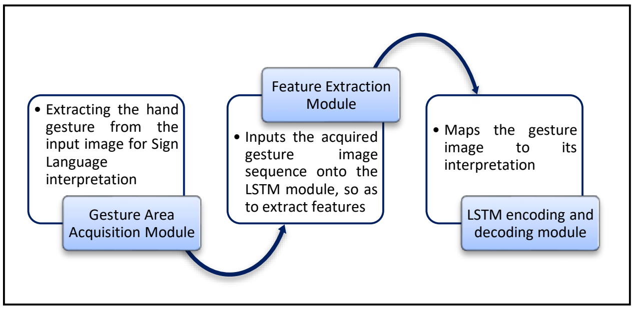


Figure 2: SLI Modules

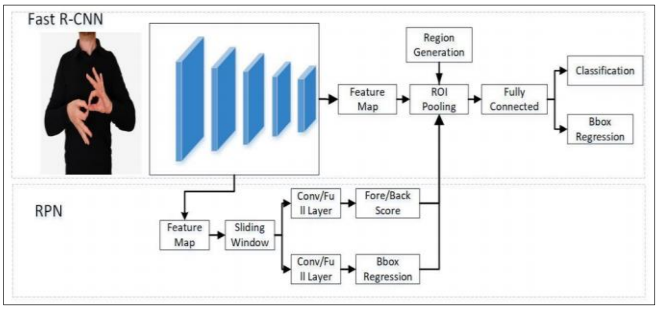


Figure 3: Structure of Faster R-CNN

### 3.2.1 Gesture Area Acquisition Module

The module uses Faster R-CNN which is the integration of the RPN (Region Proposal Network) and the Fast R-CNN module as shown in Figure 3, in order to get better performance with reference to accuracy and speed [7]. This method is used for locating hands in sign language, that can locate and track down hand gestures accurately while ignoring many of the factors like background, skin colour [1-4], hand movement or blur, etc.

### 3.2.2 Feature Extraction Module

Figure 4 depicts the 3D Convolutional Network block [7] which uses the depth sensing feature along with height and width of the hand gesture extracted to obtain precise information from the pixel. The inclusion of depth attribute helps in identifying, overlapping of the hands in the gesture image. The feature extracted by this module is loaded onto the LSTM module.

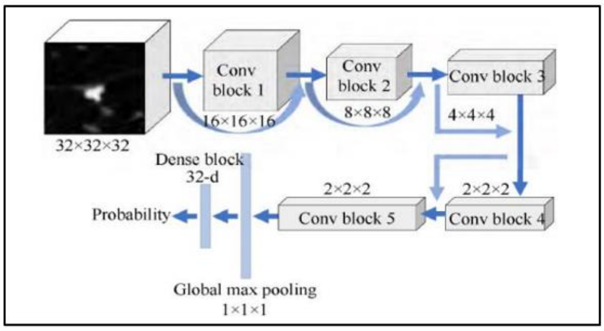


Figure 4: 3D CNN convolution blocks

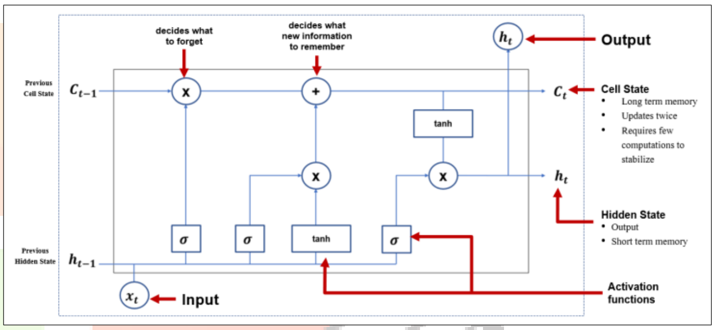


Figure 5: LSTM structure

### 3.2.3 LSTM encoding and decoding module

Long Short-Term Memory Neural Network [7-8] applies the sequence-to-sequence learning of encoding the gesture image pixels and decoding them, for sign language recognition. The module uses the information stored in the Cell State and Hidden State of the previous cell in the neural network to learn about the present Cell State and Hidden State as shown in Figure 5. The Cell State maintains the Long-term memory required for gesture identification whereas the Hidden State stores the Short-term memory used for sequence maintaining.

### 3.3 GUI and Interactions

Figure 6. gives the Wireframe for Indian SLI created using Adobe XD. The user interaction of the User and GUI are:

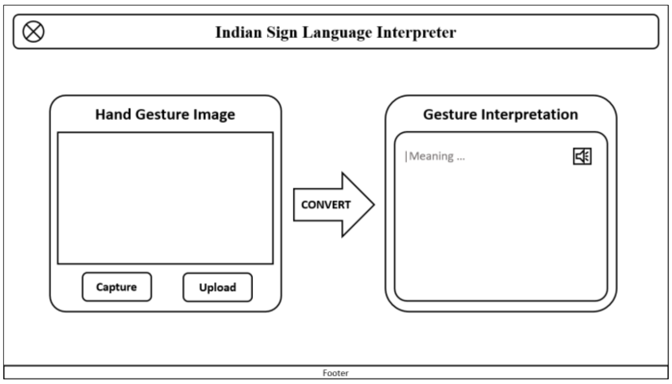


Figure 6: Web interface for SLI

- Capture: clicks a real time image of the hand sign gesture made by the signer.
- Upload: uploads an existing image (stored in local storage of the computer) of the hand sign gesture made by the signer.
- Convert: interprets the hand sign gesture, i.e., converts the image to its meaning.
- Speaker button: reads out the interpretation (text to speech).

## IV. Conclusion

The state-of-art study of various approaches to design a Sign Language Interpreter for deaf and mute people. After understanding both the approaches i.e., the Glove-Sensor based approach and the Vision-based approach used for Sign Language Interpretation as well as learning about their respective pros and cons, it was revealed that the Vision-based approach would be the most viable option as it poses no threat of hardware malfunction and yields better accuracy and efficiency results. The knowledge of the various methodologies can be used for the research of SLI. The proposed system architecture involving Gesture Area Acquisition Module, Feature Extraction Module, the LSTM Encoding and Decoding Module can be used for a better Machine Learning model.

## References


**[1]** H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.

**[2]** S. C.J. and L. A., "Signet: A Deep Learning based Indian Sign Language Recognition System," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0596-0600, doi: 10.1109/ICCSP.2019.8698006.

**[3]** L. Boppana, R. Ahamed, H. Rane and R. K. Kodali, "Assistive Sign Language Converter for Deaf and Dumb," 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 2019, pp. 302-307, doi: 10.1109.

**[4]** Apoorv, Shubhendu & Bhowmick, Sudharshan & Prabha, R., "Indian sign language interpreter using image processing and machine learning," IOP Conference Series: Materials Science and Engineering. 872. 012026. 10.1088/1757-899X/872/1/012026.

**[5]** D. Gupta, J. P. Mohanty, A. K. Swain and K. Mahapatra, "AutoGstr: Relatively Accurate Sign Language Interpreter," 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), Rourkela, India, 2019, pp. 322-323, doi: 10.1109/iSES47678.2019.00080.

**[6]** D. Gupta, J. P. Mohanty, A. K. Swain and K. Mahapatra, "AutoGstr: Relatively Accurate Sign Language Interpreter," 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), Rourkela, India, 2019, pp. 322-323, doi: 10.1109/iSES47678.2019.00080.

**[7]** S. He, "Research of a Sign Language Translation System Based on Deep Learning," 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Dublin, Ireland, 2019, pp. 392-396, doi: 10.1109/AIAM48774.2019.00083.

**[8]** K. Bantupalli and Y. Xie, "American Sign Language Recognition using Deep Learning and Computer Vision," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4896-4899, doi: 10.1109/BigData.2018.8622141.

**[9]** P. Sridevi, T. Islam, U. Debnath, N. A. Nazia, R. Chakraborty and C. Shahnaz, "Sign Language Recognition for Speech and Hearing Impaired by Image Processing in MATLAB," 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Malambe, Sri Lanka, 2018, pp. 1-4, doi: 10.1109/R10-HTC.2018.8629823.

**[10]** Shrinidhi Gindi, Amina Bhatkar, Ramsha Ansari, S Hasan Haider, "India Sign Language Translator using Residual neural network and Computer Vision," IRJET, vol 7, issue 07, July 2020

**[11]** P. C. Badhe and V. Kulkarni, "Artificial Neural Network based Indian Sign Language Recognition using hand crafted features," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225294.

**[12]** H. Rewari, V. Dixit, D. Batra and N. Hema, "Automated Sign Language Interpreter," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, 2018, pp. 1-5, doi: 10.1109/IC3.2018.8530658.

**[13]** Medet Mukushev, Arman Sabyrov, Alfarabi Imashev, "Evaluation of Manual and Non-manual Components for Sign Language Recognition," LREC, vol: Proceedings of the 12th Language Resources and Evaluation Conference, issue 16, pp 6073-6078, May 2020

**[14]** M. Agrawal, R. Ainapure, S. Agrawal, S. Bhosale and S. Desai, "Models for Hand Gesture Recognition using Deep Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2020, pp. 589-594, doi: 10.1109/ICCCA49541.2020.9250846.

**[15]** A. Er-Rady, R. Faizi, R. O. H. Thami and H. Housni, "Automatic sign language recognition: A survey," 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Fez, 2017, pp. 1-7, doi: 10.1109/ATSIP.2017.8075561.

**[16]** Z. A. Memon, M. U. Ahmed, S. T. Hussain, Z. A. Baig and U. Aziz, "Real Time Translator for Sign Languages," 2017 International Conference on Frontiers of Information Technology (FIT), Islamabad, 2017, pp. 144-148, doi: 10.1109/FIT.2017.00033.

**[17]** Jasmine Kaur and C. Rama Krishna, "An Efficient Indian Sign Language Recognition System using Sift Descriptor," 2019 International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6

**[18]** J. Ghorpade and B. Sonkamble, "Predictive Analysis of Heterogeneous Data – Techniques & Tools," 2020 5th International Conference on Computer and Communication Systems (ICCCS), Shanghai, China, 2020, pp. 40-44, doi: 10.1109/ICCCS49078.2020.9118578.

**[19]** S. MUGALA, P. JJAGWE and J. ASIIMWE, "Glove Based Sign Interpreter for Medical Emergcies," 2019 IST-Africa Week Conference (IST-Africa), Nairobi, Kenya, 2019, pp. 1-8, doi: 10.23919/ISTAFRICA.2019.8764816.

**[20]** S. A. E. El-Din and M. A. A. El-Ghany, "Sign Language Interpreter System: An alternative system for machine learning," 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 2020, pp. 332-337, doi: 10.1109/NILES50944.2020.9257958.

**[21]** M. Elmahgiubi, M. Ennajar, N. Drawil and M. S. Elbuni, "Sign language translator and gesture recognition," 2015 Global Summit on Computer & Information Technology (GSCIT), Sousse, 2015, pp. 1-6, doi: 10.1109/GSCIT.2015.7353332.