**LITERATURE REVIEW**

**Real-Time Communication System Powered by AI for**

**Specially Abled**

One of the foundations of human civilization is communication. However, substantial sections of society are not equipped with the ability to communicate easily with others. Millions of people worldwide suffer from varying degrees of hearing loss, speech impairment, or both. Sign Language is a powerful tool to help overcome these disabilities and bridge the gap between hearing and differently-abled people. There are many sign language standards across the world and with the advancement in AI and machine learning, many sign language recognition models have been deployed for a vision-based approach to detect and translate sign language into text, speech, and other formats. Some of these works relevant to this project are explained.

**[1] Indian Sign Language Recognition using Deep Learning**

P. Mistry, V. Jotaniya, P. Patel, N. Patel and M. Hasan,

*2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, 2021, *IEEE*

This paper provides a method of translation from Indian sign language to English words through deep learning models constructed using Convolutional Models and Sequence Models. The goal is to interpret a sign which is recorded through a video camera and sampled into sequences of images. This sequence is then passed through a deep learning model which passes an output vector with the highest value corresponding to the word.

The dataset is divided into training and testing purposes. The training dataset has 7 different signs (Days of week-Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday) The videos corresponding to these signs were recorded and sampled by extracting the frame sequence and developing around 50 sample sequences per word. Image augmentation was done which performed image translation, rotation, and scale to these images and increased its sample size further to about 1500 samples per word making it a 10500 sample size training dataset. Convolutional Neural Network (CNN) has been used for extracting features from images and Long Short Term Memory (LSTM) for analyzing the sequence of these features. The model converged to the training dataset and achieved an accuracy of 99.98%.

**[2] Indian Sign Language Recognition using Deep Learning**

K. Bantupalli and Y. Xie

*2018 IEEE International Conference on Big Data (Big Data), IEEE*

The paper presents a vision-based application that offers sign language translation to text thus aiding communication between signers and non-signers. The proposed model takes video sequences and extracts temporal and spatial features from them. Inception, a CNN (Convolutional Neural Network) model, is used for recognizing spatial features. An RNN (Recurrent Neural Network) is then used to train on temporal features. The gesture segments identified and processed by the CNN are classified by LSTM into one of the gesture classes using sequence data.

The dataset used is the American Sign Language Dataset. Each sign was performed five times by a single signer in varying lighting conditions and speed of signing. Each video was broken down by frame to images and then augmented to increase the data set for each sign to 2400 images. The CNN and RNN models were trained independently. The models yielded 90-93% accuracy depending upon the specifics of the architecture.

One of the problems the model faced is with facial features and skin tones. The model also suffered from loss of accuracy with the inclusion of faces, as faces of signers vary and performed poorly when there was variation in clothing.

**[3] Bengali Sign Language Recognition Using Deep Convolutional**

**Neural Network**

M. A. Hossen, A. Govindaiah, S. Sultana and A. Bhuiyan

*2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), IEEE*

In this paper, a new method for Bengali Sign Language Recognition using Deep Convolutional Neural Networks (DCNN) is proposed. The method is built to recognize static hand signs of 37 letters of the Bengali alphabet. The dataset comprises 37 hand signs (total 1147 images), in which each sign class includes 31 images taken at varying distances, orientation and intensity levels. The images represent the signs of the Bengali Language Alphabet as illustrated in the Bengali Sign Language Dictionary.

The images were augmented through many random transformations. Using the concept of transfer learning and fine-tuning, the VGG16 network was pre-trained. The pre-trained network is run on the dataset one time to extract the features necessary for the classification. Hence with the use of deep convolutional neural network and utilizing the learned features from a pre-trained network and fine-tuning the top layers of this network, a high overall recognition rate of 96.33% on the training dataset and 84.68% on the validation dataset have been achieved. This work has however used a very small dataset.