Real-Time Communication System Powered By AI For Specially Abled

Literature Survey

Paper 1

Title: SIGN LANGUAGE RECOGNITION USING IMAGE PROCESSING

Authors: Vijay More, Sanket Sangamnerkar, Vaibhav Thakare, Dnyaneshwari Mane, Rahul Dolas

Abstract:

Real time Hand Gesture Recognition and feature extraction using a web camera. In this approach, the image is captured through webcam attached to the system. First the input image is preprocessed and threshold is used to remove noise from image and smoothen the image. After this apply region filling to fill holes in the gesture or the object of interest. This helps in improving the classification and recognition step. Then select the biggest blob (biggest binary linked object) in the image and remove all small object, this is done to remove extra unwanted objects or noise from image. When the preprocessing is complete the image is passed on to feature extraction phase.The test image is classified in nearest neighbor's class in training set. The classification results are displayed to user and through the windows text to speech API gesture is translated into speech as well.

Future Work:

In this system, the data sampled from the sign language demonstration video was learned by a customized convolutional neural network. As a result, we obtained high accuracy using only 2D images obtained from a low-cost camera with much less data size than previous studies. Since the learning data used in this paper was only 20 cases, if we can obtain and use data in more various situations, we can learn more various actions with higher success rate.

Paper 2

Title: Sign Language Recognition Using Convolutional Neural Networks

Authors: Lionel Pigou , Sander Dieleman, Pieter-Jan Kindermans, and Benjamin Schrauwen

Abstract:

Recognition system using the Microsoft Kinect, convolutional neural networks (CNNs) and GPU acceleration. Instead of constructing complex handcrafted features, CNNs are able to automate the process of feature construction. We are able to recognize 20 Italian gestures with high accuracy. The predictive model is able to generalize on users and surroundings not occurring during training with a cross-validation accuracy of 91.7%.

Future Work:

The generalization capacity of CNNs in spatiotemporal data can contribute to the broader research field on automatic sign language recognition.