WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Survey Paper on Real Time Sign Language Interpreter

S. R. Aajmane[1], A.S.Neje[2], B. S. Khedkar[3], S. S. Koulage[4], S. M. Momin[5]

[1,2,3,4,5]Department of Computer science and Engineering, Dr.J.J.Magdum college of Engineering, Jaysingpur, India
[1]aajmane09@gmail.com, [2]nejeabhijeet99@gmail.com, [3]baharkhedkar25@gmail.com, [4]shreyakoulage@gmail.com, [5]sajidmomin29@gmail.com

DOI: https://doi.org/10.55248/gengpi.2022.3.5.20

ABSTRACT

Deaf and Dumb people use sign language to communicate. There are various sign recognition techniques that produce output in the form of words or identified signs. The suggested method focuses on Sign Language Interpretation in correct sentences .In addition to sign identification several NLP (Natural Language Processing) techniques are applied .Input is given as video of sign language followed by framing and segmentation on video. Deaf people get isolated as result of this. However, if an android application can be developed to convert sign language into written and audio format, the gap between normal people and the deaf community can be narrowed. The HAAR CASCADE classifier is used to identify signs. The continuous words for each sign are sent as input to the POS (Part of Speech) tagging module after sign recognition. It is utilized a word net POS tagger with its own word net dictionary. Finally, the sentence is framed using the LALR Parser. In this approach the suggested sign language Interpreter model produces intelligible sentence. This sentence is again converted into audio using gTTS API .We will be developing an android application which will scan the sign languages and interpret it into specific sentence and audio using which a normal person can understand the language of Deaf and Dumb People.

Keywords: *Sign Language, Image Recognition, Feature Extraction, Key points Creation, Hidden Morkov Model, LSTM (Long Short Term Memory), HAAR CASCADE.*

1. INTRODUCTION

The way of transferring the information from one person to another is called communication. Most of the time people use signs and words for the communication. Natural language is used by normal people to communicate/interact with each other while tactile sign language is used by deaf and dumb people to interact. Nowadays people with disabilities experience difficulties to stand in the race because of ferocious competition in every field. The effort is to develop an application which to a survey, India consists of nearly 2.4 million deaf, dumb populations which approximately make up 20% of the world's total deaf and dumb population. For hassle-free interaction between the normal person and deaf and dumb person, there is a need of an interpreter (Person who has the knowledge of sign language, as well as normal language).

Sign language is divided into two i.e., Visual Sign Language & Tactile Sign Language. a) Visual sign language: It is used by hearing & speech impaired people b) Tactile sign language: It is used by hearing & sight impaired people. We are basically working on the visual sign language used by deaf & dumb. Sign Language varies country to country it depends on its culture as Sign language in India is ISL (Indian Sign Language) America uses ASL (American Sign Language), China uses CSL (Chinese Sign Language). Sign Language is a method of communication for deaf & dumb which is composed of various gestures formed by hand shapes, body orientation & facial expression. Each gesture has a meaning assigned to it.

Alphabets in sign language are composed of different hand shapes & words are composed of hand shapes with orientation. Complete visual sign language also includes facial expressions. Visual sign language is an effective means of communication for deaf & dumb. Though it is true, the hearing-impaired have to challenge communication obstacles in a mostly hearing capable society. This research work will concentrate on Visual Sign Language interaction. Natural language is a skill used for understanding human language. It is a part of linguistics and Artificial Intelligence.

NLP is a step for developing a system that can convert the text (words) in human language. POS tagging is the method of NLP and first introduced in 1960. It is an important method for language processing. For many NLP applications it is the simplest and most stable step. Part of Speech tagging is the initial step for machine translation, retrieval of information and etc. Second important method in NLP is parsing. Parsing is the method which is followed by the compiler.

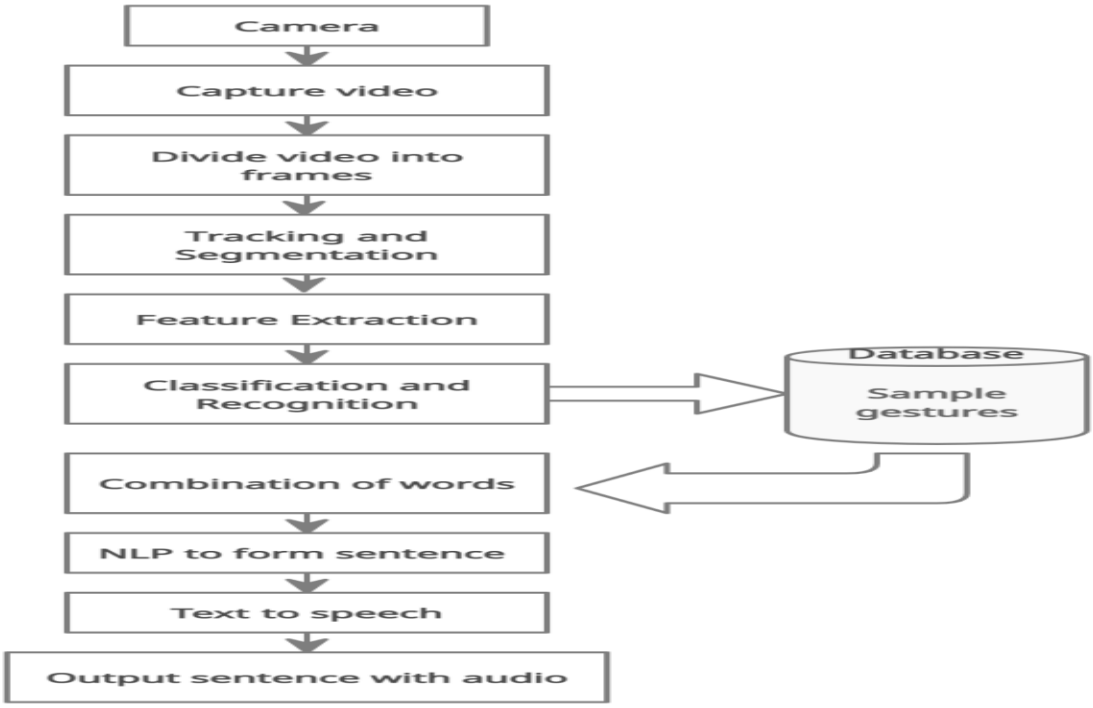

Figure 1. System Architecture of Real Time Sign language Interpreter

The signs made by deaf or dumb person are detected through camera of the user's device. It is in video format containing some number of frames, from these frames the unwanted frames are removed and useful frames are tracked and extracted. Now using MediaPipe Holistic model the posture, hand gestures and face is detected and key points are plotted. These key points are then sent to the A.I model to predict a particular sign. Signs will just represent words, so using Natural Language Processing meaningful sentences are formed using these words.

Modules:

1. Video to frames conversion
2. Key points Creation
3. Creating Dataset
4. Training a CNN on the captured Dataset
5. Predicting the Gesture

2. LITERATURE REVIEW

1. Review Of The Application Of Artificial Intelligence In Sign Language Recognition System-July 2020

This research uses Artificial Intelligence Natural Language Processing Gesture Sign Language Human communication is the process through which people comprehend what is spoken to them and express their thoughts, desires, and feelings to others, primarily through speech. When it comes to persons who have hearing loss, however, sign language is unavoidable. As a result, sign language is the most natural and effective means for deaf and hearing individuals to communicate. [1]

This research looked at a variety of AI-based systems for recognizing and interpreting sign language for machine and human comprehension. Several researchers designed and implemented a variety of computer and Android-based applications to facilitate learning and teaching sign language, while others developed a variety of machine learning techniques and frameworks to support sign language recognition systems, according to this study.

This research found that artificial intelligence has substantially improved the teaching, learning, and communication of sign language, and that it would, undoubtedly, be capable of resolving future issues in this area.

2. Arab Sign Language Recognition System -2019

Proposed an arab sign language recognition system based on CNN, inspired from LeNet-5 . Dataset contained 7869 images of arab signs of numbers and letters. Various experiments were conducted by varying the number of training sets from 50% to 80%. 90% accuracy was obtained with 80% training dataset. The author has also compared the results obtained with machine learning algorithms like KNN (k-nearest neighbor) and SVM (support vector machine) to show performance of the system. This model was purely image based and it can be extended to video based recognition.[2]

3. American Sign Language Recognition System -2018

Kshitij Bantupalli and Ying Xie developed an American gesture recognition system that uses CNN, LSTM, and RNN to detect video sequences. Inception, a CNN model, was used to extract spatial characteristics from frames, while LSTM was used to extract longer time dependencies and RNN was used to recover temporal features. Various trials were carried out with different sample sizes, and the dataset encompassed 100 different signs done by 5 signers, with a high precision of 93 percent. For longer temporal dependencies, the sequence is passed into an LSTM. To extract temporal features from the softmax layer, the outputs of the softmax layer and the maximum pooling layer are input into an RNN architecture [3].

4. Sign Language Recognition System for Deaf and Dumb People using Image Processing-March-2016.

This research uses Sign language identification, Hidden Morkov Model, Artificial Neural Network, Data glove, Leap motion controller, Kinetic Sensor. Communication between a deaf-mute and a hearing person has always been difficult [4].

This research examines various approaches to lowering communication barriers by building an assistive gadget for deaf-mute people. The evolution of embedded systems allows for the creation and development of a sign language translation system to assist the deaf. There are a variety of assistant tools available. The major goal is to create a real-time integrated technology that will help physically challenged people communicate more effectively.

5. Sign Language Recognition for Deaf and Dumb People-March 2015

This research uses Sign Language, Feature Extraction, Sign Recognition; PCA Gestures are a type of physical activity that a person uses to communicate important information. Gestures are a strong form of human communication. Sign language is extremely important for persons who are deaf or mute and have difficulty hearing or speaking. It is the only means of communication for such people, therefore it is critical that others comprehend their language.[5]

This project's goal is to create a system that will assist deaf and hard of hearing persons in communicating with others. It accurately and quickly recognizes hand motions and transforms them into voice so that individuals who aren't familiar with sign language may comprehend what's going on. It employs Eigen Value Weighted Euclidean Distance Based Classification Technique and is built in MATLAB. The sound associated with the recognized symbol is produced as the output.

3. ALGORITHM'S

1. Long Short Term Memory:

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This is a behavior required in complex problem domains like machine translation, speech recognition, and more.LSTMs are a complex area of deep learning.

2. Reinforcement Neural Network :

Reinforcement learning is a goal-directed computational approach where an agent learns to perform a task by interacting with an unknown dynamic environment. During training, the learning algorithm updates the agent policy parameters. The goal of the learning algorithm is to find an optimal policy that maximizes the long-term reward received during the task. Depending on the type of agent, the policy is represented by one or more policy and value function representations. You can implement these representations using deep neural networks. You can then train these networks using Reinforcement Learning Toolbox software.

3. Convolution Neural Network:

CNNs are regularized versions of multilayer perceptron's. Multilayer perceptron's usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them prone to over fitting data. Typical ways of regularization, or preventing over fitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

4. Artificial Neural Network:

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron receives signals then processes

them and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called edges. Neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traveling the layers multiple times.

5. Hidden Markov Model

A Markov Chain is a model or a type of random process that explains the probabilities of sequences of random variables, commonly known as states. Each of the states can take values from some set. In other words, we can explain it as the probability of being in a state, which depends on the previous state. We use the Markov Chain when we need to calculate the probability for a sequence of observable events. However, in most cases, the chain is hidden or invisible, and each state randomly generates 1 out of every *k* observations visible to us. Now, we will define the Hidden Markov Model.

4. MODULES

1. Video to frames conversion:

Using OpenCv the input is taken as an video this video will be then broken down into 30 frames and these 30 images or frames will be then used for plotting and predicting the sign.

2. Key points Creation:

After getting input from the user, these images are then passed to a function which plots key points and their connection using mediapipe holistic model and these key points are drawn on the screen for the user to see using matplotlib library. Points indicates joints of body and lines indicates connection between joints.

3. Creating Datasets:

We will be having a live feed from the web-cam and every frame that detects a gesture is created and keypoint values will be saved in terms of numpy array. These numpy array values are saved for different signs. Each sign will have 30 different types of numpy arrays.

4. Training a LSTM on the captured dataset:

We now train a LSTM on the newly produced dataset. To begin, we load the data using keras, We will be developing the datasets ourselves for this project because we need datasets in terms of mediapipe holistic keypoint values which is not available as per our requirements on the internet. Every frame that detects a posture will be converted into numpy array and saved in a directory that comprises folders with the name of signs, each containing 30 files acquired during dataset creation.

We will then train and test the model using the flow from directory function to import the train and test set data, with the names of the number folders serving as the class names for the images loaded.

5. Predicting the Gesture:

After successfully training the model with the given dataset. We now directly convert the live frames from the camera into numpy arrays and pass them to the model as input, Now the model will predict based on the keypoint values. The sentence having maximum accuracy with minimum loss will be shown to the user in terms of textbox.

5. CONCLUSIONS

A survey on sign language recognition is offered in this study, and several strategies for the same have been investigated and analyzed. In the recognition process, segmentation is important because it separates the skin region from the backdrop, which influences recognition accuracy. Apart from segmentation, classification relies on feature extraction techniques, which minimize dimensionality and lower computation costs. Deep neural networks (CNN, Inception model, LSTM) outperform classic classifiers like KNN and SVM, according to a study of various classification techniques. Furthermore, many papers require hardware, which may raise costs, and while everyone has developed algorithm, there is no application to interact with deaf, stupid, or normal people.

REFERENCES

[1] Review Of The Application Of Artificial Intelligence In Sign Language Recognition System-July 2020. Publisheb by Oluwashina Akinloye Oyeniran, Kehinde Sotonwa, Joshua Oyeniyi .

[2] S. Hayani, M. Benaddy, O. El Meslouhi and M. Kardouchi, "Arab Sign language Recognition with Convolutional Neural Networks," 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), Agadir, Morocco, 2019, pp. 1-4, doi: 10.1109/ICCSRE.2019.8807586.

[3] 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.

[4] Sign Language Recognition System For Deaf And Dumb People using Image Processing-March-2016 Published by Manisha U.Kakde , Mahender G. Nakrani2 , Amit M. Rawate3.

[5] Sign Language Recognition for Deaf and Dumb People-March 2015 Publisheb by Y.M.Pathan1, S.R.Waghmare2, P.K.Patil.