**INTRODUCTION**

A Sign Language is a language in which communication between people is made by visually transmitting the sign patterns to express the meaning. It is a replacement of speech for hearing and speech impaired people. Thus, because of which has attracted many researchers in this field from long. Many researchers have been working in different sign languages like American Sign Language, British Sign Language, Taiwanese Sign Language, etc. but few works has been made progress on Indian Sign Language.

The hearing impaired people become neglected from the society because the normal people never try to learn ISL nor try to interact with the hearing impaired people. This becomes a curse for them and so they mostly remain uneducated and isolated. Thus recognition of sign language was introduced which has not only been important from engineering point of view but also for the impact on society.

Studies prove that Children feel more lonely when they are silent and the problem of this communication gap will add to their depression. There are also a numerous number of problems prevailing due to this communication gap, our project aims to act as a bridge to close this distance and make the flow of thoughts more easy.

**PROBLEM STATEMENT**

Even with all the advancements in Artificial Intelligence and machine learning , there is only a little progress in sign language recognition. But with recent developments in computer vision and a massive rise in computational power we can solve problems that require intensive processing.

Our Aim is to recognise the signs and gestures of Indian sign language and convert them into world wide understandable language.

**ABSTRACT**

There are over 450 million sign language users all over the world. There are many sign languages used across the world. They are identified and named according to the geographical location like

1. Indian sign language (used in india)
2. American sign language (used in america)
3. British sign language (used in britain)

As we know there is no problem for mute people while Communicating among themselves. But a problem comes when mute people try to communicate with unmute people who cannot understand. So we suggest a solution using a machine learning algorithm to recognize the signs and gestures of Indian sign language and convert them into a world wide understandable language. (English)

Our project explores different methods and algorithms to recognize the gestures from indian sign language from live video and convert into a proper english sentence. We use different algorithms , analyse their accuracies and tweak parameters to increase the efficiency and also use NLP algorithms to generate proper English sentences.

**EXISTING SYSTEM**

There are a few already existing solutions to recognise the signs in sign language

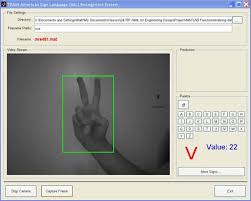
Like

1. **Sensor gloves**

Sensor gloves are synthetic gloves with in-built sensors which are able to detect the position of the hands, fingers and knuckles. The position of the hands, dingers and knuckles are plotted in 3d obtaining a configuration. these configurations have already defined signs and converted accordingly.

They are highly accurate because of their static lookup nature. The system directly maps the identified configuration with already configured gesture. They are easy to use i.e., no additional training required to use these gloves , once after wearing the gloves the user can use gestures like he would regularly do without any gloves. The implementation implicitly does not involve any machine learning algorithms.

The use of physical hardware makes the system expensive and as a result a very small population will be able to afford it making it highly not feasible. Carrying gloves wherever you go, wearing it all the time makes users fell conscious and uncomfortable and also considering only the hand gestures gives no room to get the context of the expression.

1. **Software Algorithms**

Software Algorithms collectively refers to the machine learning and artificial

algorithms using computer vision and gesture recognition to translate sign

language. Different models are trained on different instances of datasets enabling

the algorithms to be able to recognise the gesture in sign language.

Due to low implementation cost they are usually highly affordable and

economical. Can offer high accuracy if designed and trained correctly. They can

be easily deployed and scaled making them available and reliable.The sign space

being very large results in considerably long training times. Achieving a high

accuracy rate may be difficult due to a major part of the system being uncertain.

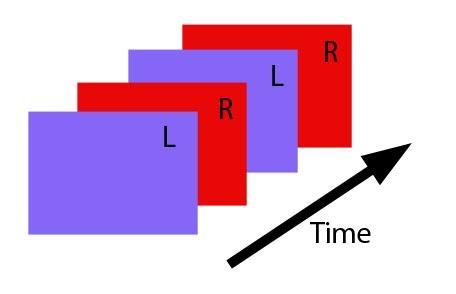
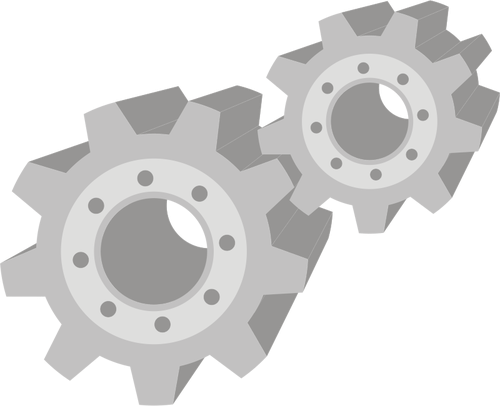
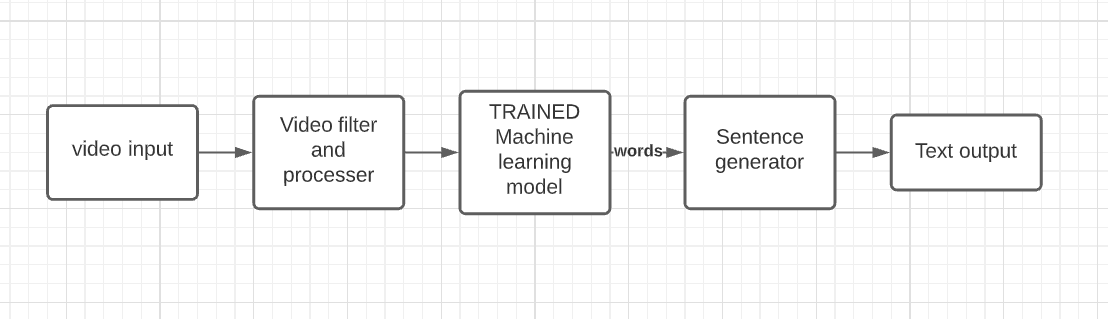
**PROPOSED SYSTEM**

Keeping in mind all the advancements in the mind it is not so hard to solve this problem using machine learning and artificial intelligence . so, we propose a solution to this problem using machine learning .

* We train a Machine learning model using supervised learning on the signs and gestures in Indian sign language so that it can identify the signs and gestures in the Indian sign language in future.
* This model takes the input as an area of interest generated by the object identifier and finds their corresponding English words.
* The generated words are then passed to a sentence generator at which the generated words are framed into a proper english sentence.

**ARCHITECTURE**

The Video Input is taken as input and each frame is taken out of the video. The Video filter and processor labels the hand gestures in the frames. These labelled frames are then sent to the learning model. This model now gets trained by the sent data. After the gestures are recognised, they’re sent to the sentence generator. The Sentence generator takes the text and phrases the sentence, words accordingly for understandable output to the user. This output of the Sentence generator is shown as Output which is an understandable sentence.



**Video input**

Recorded video in real time which is captured from camera from the computer system web cam. It takes the video and generated frames and passes them to the video filter.

**Video filter and processor**

The video filter takes the generated frames and identifies the objects and extracts the areas of interest from each frame . We use object detection algorithms to identify hands and faces.

**Trained machine learning model**

A trained machine learning model that can recognise the signs from the areas of interest and also predict the context based on the facial expressions.The model generates english words corresponding to the signs in the video.

**Sentence generator**

The sentence generator is a trained natural language processing algorithm (NLP) that is used to make properly framed English sentences. It takes generated words from the ML model as input and uses this as key words to generate English sentences.

**Text Output**

The text output generated from the sentence generator is shown as the output and read outloud from the device speaker.

**FUNCTIONAL REQUIREMENTS**

1. **Sign language recognition**

Our system must be able to recognise the gestures in Indian sign language. The input is in the form of video captured containing Indian signs and we must recognise the signs in the video and translate them to English.

We achieve this using a trained machine learning model.We train the model on the datasets containing Indian sign language gestures so that it can identify the Indian sign language signs in the video.

**Live video**

The system must be capable of capturing live video containing the Indian sign language signs and converting into English.

**Recorded video**

The system must be able to take a recorded video which contains indian sign language signs as an input and translate into english.

1. **Sentence generator**

The Sign language recognition system recognises the signs in the video and generates the corresponding words to the signs in English.

The sentence generator must take these generated words as input and Frame grammatically correct, understandable and meaningful English sentences.

1. **Object recognition**

The identifications , recognition and translations are dependent on the Hand movements and face gestures of the user. From the video captured the system must be able to successfully identify and extract the hands and face of the user. We can do this using object recognition.

1. **Reading out loud**

The output generated must be played using divide audio enabling the vision imapaired people to know what the person is saying.

**NON-FUNCTIONAL REQUIREMENTS**

1. **Accuracy**

The accuracy of the system means that the correctness of the recognition of the gestures in the video and converting them to English.

As this has to be a real time system the accuracy must be very high and there should be no place for errors.

1. **Reliability**

The users of the system are dependent on the system for the conversion of the Indian sign language to English . So , in addition to the Accuracy we must also provide high availability and dependability.

1. **Processing time and cost**

Since the translation is a computationally intensive task, the processing time may be high but this must be reduced as much as possible and the cost of processing must not be borne by the user.

1. **Sign space**

There are nearly 10,000 signs in Indian sign language and to Implement a Real Time translation system there must be a considerable amount of these words included in the training phases of the system.

There are also some regional sign changes due to the large diversity of India. There may be different versions of Indian sign language according to the geographical region of the Indian subcontinent. Keeping these irregularities aside we Included 100 most commonly used Indian sign language gestures. If we can get high accuracy for these 100 signs the remaining can be trained gradually using this algorithim.

1. **Responsive**

This translation system is a computationally intensive system even so , the responsive nature of the application must not be affected by this. The GUi must be Independent and responsive irrespective of the Computational Intensity of the generated sentence.

**LITERATURE SURVEY**

This systems demands for huge number of technologies and exploring different algorithms the technologies and algorithms are listed below

**Machine learning**

Machine learning (ML) is the study of computer algorithms that can improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so.

**Artificial Intelligence**

artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience.

**Natural Language Processing**

Natural language processing (NLP) refers to the branch of computer science—and more specifically, the branch of artificial intelligence—concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.

There is no currently working Sign language translation system. So, it is not very easy to develop the algorithm and improve the accuracy in a short time. The already implemented systems are either less accurate or has low sign space and also no high level conversion.

**Challenges For Solving the problem**

1. No dataset available

There is no standard dataset available in video format to train the system.

1. No feasible model till date

There is no standard Algorithm to train using video data and identify the signs and gestures from video data.

**Algorithms Explored**

1. **Convolution Neural Networks**

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

1. **Optical Flow**

Optical flow provides a concise description of both the regions of the image undergoing motion and the velocity of motion. In practice, computation of optical flow is susceptible to noise and illumination changes.

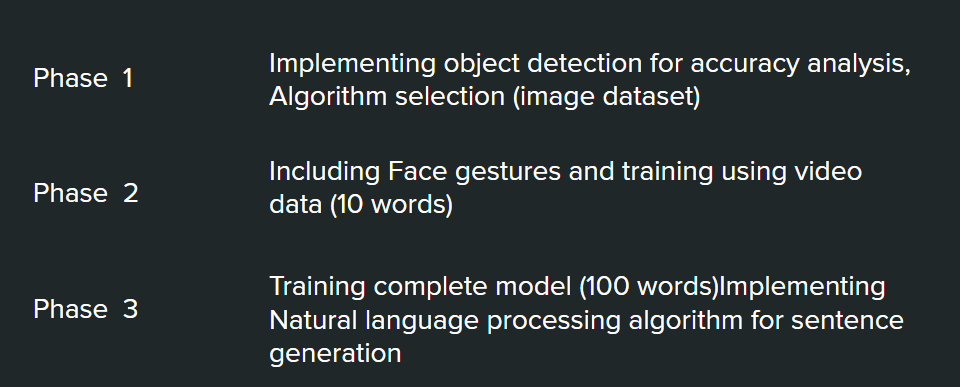
1. **Key frame Extraction**

Key frames are the frames that alone can define the content of the video. We follow a map-like structure and identify the gesture based on the extracted key frames.

**How to overcome these Challenges**

1. There is no standard dataset available so we decided to make our own dataset
2. We are also including the face gestures so that we can also keep the context  
    Of the sentence.
3. Decomposition of the system is important

**Decomposed system workflow**



**Phase 1**

We take image dataset and train the model because working with video data is tedious.we can fix on the algorithm and parameters in this phase.

Dataset in phase1

Single hand gestures <https://www.kaggle.com/muhammadkhalid/sign-language-for-alphabets>

Double hand gestures <https://www.kaggle.com/vaishnaviasonawane/indian-sign-language-dataset>

**Phase 2**

We now train the model with video data using 10 gestures initially by using the algorithm fixed in phase 1. After fixing the algorithm we also add the face gestures.

**Phase 3**

Train the model with complete 100 gestures after getting high accuracy we then need to work on Sentence generator algorithm

**CONCLUSION**

The Sign Language Translation system can change the lives of many people across the world. It can be used to talk to the mute people and understand them better. This will help to communicate with children at the younger ages well and therefore help them to get out of their loneliness and depression. In Adults this translation system opens up more opportunities to those who are backward until now due to the communication gap.

The development of this system helps the people using Indian sign language and also using this model we can change the training dataset and implement translators for any sign language. Also we can use transfer learning to increase the sign space for recognition. Adding sign generation for English to this system will lead to an even more decrease in communication gap.

**REFERENCES**

1. Accuracy analysis paper

<https://www.sciencedirect.com/science/article/pii/S1877050917320720>

1. Hardware Implementation paper

<https://www.sciencedirect.com/science/article/pii/S1877050918321331>

1. Software implementation paper

<https://drive.google.com/file/d/1A37c4Ta9CVVFyhwjIFb3Uw_PJCsKs0lb/view?usp=sharing>

1. Other papers

<https://drive.google.com/file/d/1gOO20k-eyH7GtNm1PfjSQ5wXD9w6_5kD/view?usp=sharing>

<https://drive.google.com/file/d/1EIQ-XWRxuzSEGoclHD1yBYRLzSeUpB18/view?usp=sharing>

<https://drive.google.com/file/d/14sdlBv7UYz2qbprNgnMNL1QyvZBZFyXU/view?usp=sharing>

1. Selected words

<https://docs.google.com/document/d/1uq2HevDOJ6iAcZ3Sb9WJ8338SDogyoN1sAAxGpcL6tM/edit?usp=sharing>

1. CNN

<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

1. Optical flow

<https://www.sciencedirect.com/topics/engineering/optical-flow>

1. Official Indian Sign language reference

<https://indiansignlanguage.org/>

<http://www.islrtc.nic.in/>