A Project Identification Report

On

Indian Sign Language Recognition

Submitted by:

Team Infinity (AICB1 – AI06)

01	**Hari Lakshmi A. S**
02	**Anushka Bajirao Patil**
03	**Mekapothu Manikanta Reddy**
04	**Pavaneshwar Akula**
05	**Vrushali Arun Ghauge**
06	**Sikandar Baksh**

Under the guidance of

Ms. Diana Baby

Project guide cum Mentor

edunet foundation

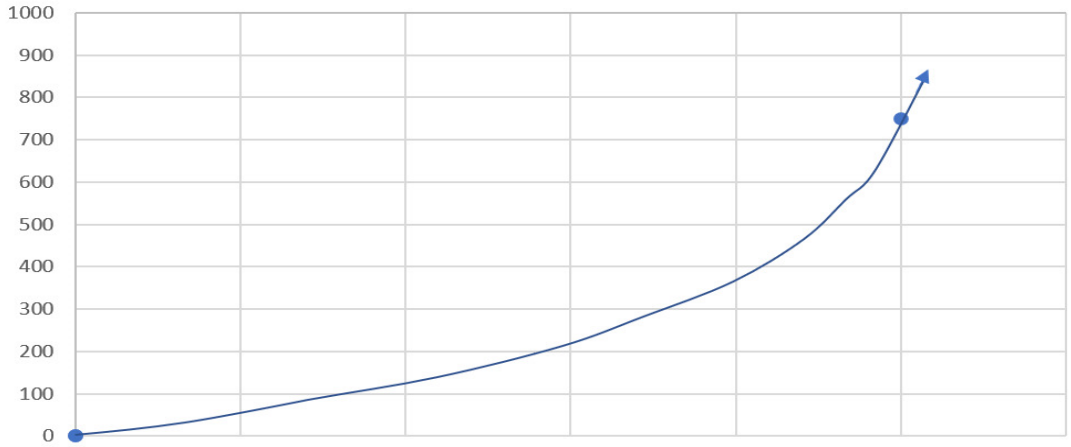
BACKGROUND OF THE PROJECT:

Communication is very crucial to human beings, as it enables us to express ourselves. We communicate through speech, gestures, body language, reading, writing, or through visual aids, speech being one of the most commonly used among them. However, unfortunately, for the speaking and hearing-impaired minority, there is a communication gap. Visual aids, or an interpreter, are used for communicating with them. However, these methods are rather cumbersome and expensive, and can't be used in an emergency.

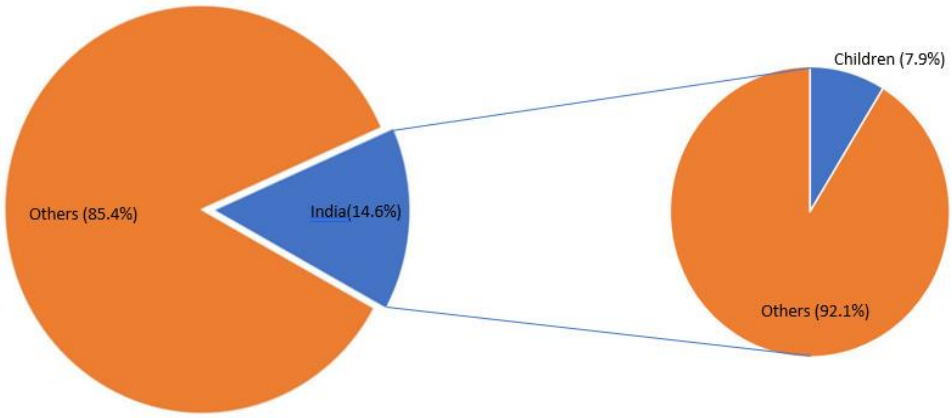
Around 466 million people worldwide have hearing loss, and 34 million are children. `Deaf' people have very little or no hearing ability. They use sign language for communication. People use different sign languages in different parts of the world. Compared to spoken languages they are very less in number. India has developed a sign language by the name Indian Sign Language (ISL). In developing countries, there are only countable schools for deaf students. The Unemployment rate among adults with hearing loss is very high in developing countries. Data from Ethnologue states that among the deaf population in India, about 1% of the total population, the literacy rate and the number of children attending schools are very low. Our project aims to reduce the basic step in bridging the communication gap between normal people and deaf people using Indian sign language. Effective extension of this project to words and common expressions may not only make deaf people communicate faster and easier with the outer world, but also provide a boost in Developing autonomous systems for understanding and aiding them.

USERS / CUSTOMERS:

1. *Schools and Institutes that train such differently-abled people* – According to a report in Hindustan Times, the Delhi High Court issued a notice listing a PIL filed by Nipun Malhotra seeking official recognition of ISL and the government appears to wake up to this. On 23 September 2021, the minister of states for social justice and empowerment said he would write to home minister Rajnath Singh that there should be organizations that train on Indian Sign Language (ISL), as a result of this 700+ schools will be starting ISL by 2025.



2. *Deaf and Dumb people* - According to a WHO report, more than 14% of the deaf population of the world is in India out of which ~8% of them are children.



Source: World Health Organization Report

PROBLEM STATEMENT:

Sign language uses lots of gestures so that it looks like movement language which consists of a series of hand and arm motions. sign language includes specific gestures to each alphabet in the English dictionary and for each number between 0 and 9. Based on these sign languages are made up of two groups, namely static gesture, and dynamic gesture. The static gesture is used for alphabet and number representation, whereas the dynamic gesture is used for specific concepts. Dynamic also includes words, sentences, etc. The static gesture consists of hand gestures, whereas the latter includes the motion of hands, head, or both. Sign language is a visual language and consists of 3 majors

components, such as finger-spelling, word-level sign vocabulary, and non-manual features. Finger-spelling is used to spell words letter by letter and convey the message whereas the latter is keyword-based. But the design of a sign language translator is quite challenging despite many research efforts during the last few decades. Also, even the same signs have significantly different appearances for different signers and different viewpoints.

SOLUTION:

Gesture Recognition (GR) is one of the most important sub-topics in action recognition. In recent years, it has gained much attention for its role in human-machine interaction. Gesture recognition is also used in sign language recognition which is very important for special people. A lot of research had been done in GR.

We will try to make an AI learning model which will store some data (given at the time of training) and then the module will be trained by giving so many pictorial data, here multiple data will be stored for one gesture (15-25) so that the module can detect the exact sign shown to it from any angle.

The training will be done by acting gestures manually and storing the dataset with a particular name title.

JOURNAL PAPERS:

Journal paper 1:

Referred Journal paper: Sign Language prediction using Machine learning algorithm published by International Research Journal of Engineering and Technology (IRJET). The authors who contributed to the journal are Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, and Ms. Mayuri A. Khandge.

BACKGROUND OF THE PROBLEM:

Sign Language is one of the ways to communicate with deaf people. One should learn sign language to interact with them. Learning usually takes place in peer groups. There are very few study materials available for sign learning.

Because of this, the process of learning sign language learning is a very difficult task. Most of the existing tools for sign language learning use external sensors which are costly.

PROJECT APPROACH:

In this research paper, they have used fourfold cross-validated results for the different approaches. They used Computer vision and Machine Learning algorithms for their research.

CHALLENGES:

1. The sign language is American Sign Language (ASL), British Sign Language, and Chinese Sign language, each country has its language. Extensive work has been done on American sign language recognition but Indian sign language (ISL) differs significantly from ASL. ISL uses two hands for communicating whereas ASL uses a single hand to communicate. Using both hands often leads to the security of features duet overlapping of hands.

2. There are many datasets available for ASL whereas a small number of the dataset is available for ISL.

3. Unlike ASL, Our ISL uses both hands for making gestures which leads to occlusion of the feature.

METHODOLOGY:

They classified their approach to tackle the classification problem into three stages.

Stage 1: The first stage is to segment the skin part from the image

Stage 2: Is to extract relevant features from the skin segmented images which can prove significant for the next stage.

Stage 3: Use the extracted features as input into various supervised learning models for training and then finally use the trained models for classification.

DATASET:

The dataset used in this research paper is the skin segmentation dataset from UCI containing about 2,00,000 points.

FEATURE EXTRACTION:

1. They are used to start SIFT (Scale Inverse Feature Transform) features as it computes the key points in the image.

2. The skin segmented images were obtained using the YUV-YIQ model

ALGORITHMS:

1. Support Vector Machine: The best accuracies observed for this algorithm will be 4.76% on HOG feature vectors, whereas the 'RBF' kernel failed miserably

2. Random Forest: They used Random Forest with HOG feature vectors on the compressed images and got an accuracy of 46.45 % 4-fold accuracy.

3. Hierarchical Classification: They First trained a linear kernel SVM model to classify alphabets as one-handed or two-handed. The model accuracy was 95% a. Then they trained linear kernel, Multiclass SVM models, to classify the one-handed alphabets (56% accuracy) and two-handed alphabets (60% accuracy) and combined the system. Even though the individual models performed better than the direct multiclass SVM on HOG features, overall, the performance was nearly the same and four-fold CV accuracy of 53.23% was observed.

PROJECT WORKFLOW:

1. Image capture: According to the authors, the selection of a good webcam & its interfacing is an important task of this method.

2. Image pre-processing: Image pre-processing contains cropping, filtering, brightness & contrast adjustment & many more. To do such a process the authors used Image enhancement, Image cropping & Image Segmentation methods.

Captured Images are in the form of RGB. So, they converted it to binary images and then cropped the image to get rid of the unwanted part. Then they proceeded with enhancement in certain selected areas. In Image segmentation, the Edge detection method is used which can detect the boundary of cropped.

4. Feature Extraction: There are mainly two types of feature extraction methods involved in sign recognition, first is Contour-based shape representation and description methods & another is Region-based shape representation and description methods. In this proposed method 7Hu moments technique is used & from that 7 moments are found.

RESULTS AND OBSERVATION:

They divided the data set divided into two groups, one used for training and the other for testing. The training set consists of 70% of the aggregate data and the remaining 30% is used as testing. They also performed experiments on the same (30% or 70%)

A dataset that is training as well as testing for the KNN classifier. The results of these experiments had a 100% accuracy rate.

Journal paper 2:

Referred Journal paper: Real-time Indian Sign Language (ISL) Recognition published by Institute of Electrical and Electronics Engineers. The authors who contributed to the journal are Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar

BACKGROUND OF THE PROBLEM:

This system attempts to bridge the communication gap between the hearing and speech impaired and the rest of society. The existing solutions either provide relatively low accuracy or do not work in real-time

PROJECT APPROACH:

The author modeled in a way such that Sign Language is captured from a smartphone camera and its frames are transmitted to a remote server for processing. They avoided the use of any external hardware and made it user-

friendly. Techniques such as Face detection, Object stabilization, and Skin color Segmentation are used.

for hand detection and tracking. The image is further subjected to a Grid-based Feature Extraction technique. Hand poses are then classified using the k-Nearest Neighbour's algorithm. On the other hand, for gesture classification, they fed observation sequences are fed to Hidden Markov Model chains. Using this methodology, the system achieved an accuracy of 99.7% for static hand poses and an accuracy of 97.23% for gesture recognition.

IMPLEMENTATION:

Using an Android Smartphone, gestures and signs performed by the person using ISL are captured and their frames are transmitted to the server for processing. To make the frames ready for recognition of gestures and hand poses, they need to be pre-processed. The pre-processing first involves face removal, stabilization, and skin color segmentation to remove background details and later morphology operations to reduce noise. Then the images are fed into the classifier. This is encoded for HMM and fed to it. The gesture whose HMM chain gives the highest score with the forward-backward algorithm is determined to be the recognized gesture for this pattern.

DATASET:

The authors used a dataset that consists of 0 to 9 Sign language representation, A to Z alphabets representation, and some basic words like good morning, good night, etc. In total the dataset consisted of 24624 images.

PROJECT WORKFLOW:

1. Preparation of dataset
2. Pre-processing
    1. Face Detection and elimination
    2. Skin color segmentation

3. Morphology operations
4. Object Stabilization using facial reference

3. Hand extraction and tracking
4. Feature extraction using Grid-based fragmentation technique
5. Classification
    1. Recognition of ISL Hand poses using K-NN
    2. Gesture Classification using HMM
    3. Temporal Segmentation

EXPERIMENTAL RESULTS:

They applied 6 different grid sizes – 5x5 10x10, 10x15, 15x15, 15x20 and 20x20 to extract features from the same training data. These features were then fitted into a k-NN classifier. The features extracted from the testing data were then classified using the k-NN classifier trained previously. They inferred that the grid size of 10x10 gives the highest accuracy of 99.714%.

APPLICATION:

The system is implemented as an Android application. The application uses the Smartphone's camera to capture the sign language used by the person. The frames were captured at a rate of 5 frames per second and they were connected to a remote server. After each pose or gesture is classified, the result is sent back to the application which is displayed in the top portion. They used Sockets to stimulate a client-server connection.

Journal paper 3:

Referred Journal paper: Conversion of Sign Language into Text published by. International Journal of Applied Engineering Research. The authors who contributed to the journal are Mahesh Kumar N

PROJECT APPROACH:

Sign Language is mainly used for communication with deaf-dumb people. In this paper, they used the sign language recognition of 26 hand gestures in Indian sign language using MATLAB. In this paper, they proposed a system that contains four modules such as pre-processing and hand segmentation, feature extraction, sign recognition, and sign-to-text. They used Eigen values and Eigen vectors which are used in recognition. The Linear Discriminate Analysis (LDA) algorithm was used for gesture recognition and recognized gestures are converted into text and voice format.

METHODOLOGY:

Sign Language Recognition System has different approaches Glove based approaches and vision-based approaches

1. Glove-based Approach: This approach needs the user to wear a sensor glove or a colored glove. The task will be simplified during the segmentation process by wearing a glove. The drawback of this approach is the user needs to wear gloves every time.

2. vision-based approach: There are again two different approaches to vision-based sign language recognition:

1.3D model-based

2. Appearance-based

In this paper, they used LDA Algorithm for sign Recognition. LDA is mainly used in statistics, pattern recognition, and machine learning. It is used to find a linear combination of features that characterizes or separates two or more classes of objects or events. The LDA and FLD are used as linear classifiers.

PROJECT APPROACH:

Data Acquisition ⟹ Pre-processing ⟹ Image acquisition ⟹ ⟹ Segmentation

MORPHOLOGICAL FILTERING:

They used Morphological Filtering tools which are useful for the representation and description of shape. They used the Dilation and Erosion method, where Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries

FEATURE EXTRACTION:

The reduction of data dimensionality by encoding related information in a compressed representation and removing less discriminative data is called as Feature Extraction Technique. They used their principal component is used as the main features. The following figure 7 explains the feature extraction method

The gesture recognition using the LDA algorithm that involves two phases

- Training Phase
- Recognition Phase

Here, each gesture is represented as a column vector in the training phase. Next, the algorithm finds the eigenvectors of the covariance matrix of normalized gestures by using a speed-up technique that reduces the number of multiplications to be performed.

In the recognition phase, a subject gesture is normalized concerning the average gesture and then projected onto gesture space using the eigenvector matrix. Lastly, Euclidean distance is computed between this projection and all known projections and the minimum value is selected.

CONCLUSION:

By using the LDA algorithm for sign recognition operation the dimensionality will be reduced. Due to dimensionality reduction, the noise will be reduced and with high accuracy. In the future, this project will be enhanced by determining the numbers which will be shown in words.

DATASET:

- In our project we are going to deal with the Indian Sign language dataset (ISL).

- We found this dataset in Cagle. The mentioned dataset contains Alphabetic Indian sign language representation (A TO Z) and Numeric Indian sign representation (0 TO 9). The dataset contains 42,700 files each containing an ‘n’ number of images sufficient to train the model.

- In addition to that our team will develop the dataset by adding some greeting words like Hello, Thank you, etc.