**SIGN LANGUAGE DETECTION**

**A Project Work Synopsis**

*Submitted in the partial fulfilment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Submitted by:**

20BCS6760 HRITVIK MATHUR

20BCS6712 AMISHA KHANNA

**Under the Supervision of:**

**SIDDHARTH KUMAR**



**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,**

**PUNJAB**

**March, 2022**

# Abstract

Research in the field of sign language recognition has made significant advances in recent years. The present achievements provide the basis for future applications with the objective of supporting the integration of deaf people into the hearing society. Translation systems, for example, could facilitate communication between deaf and hearing people in public situations. Further applications, such as user interfaces and automatic indexing of signed videos, become feasible. The current state in sign language recognition is roughly 30 years behind speech recognition, which corresponds to the gradual transition from isolated to continuous recognition for small vocabulary tasks. Research efforts were mainly focused on robust feature extraction or statistical modelling of signs. However, current recognition systems are still designed for signer-dependent operation under laboratory conditions. This paper describes a comprehensive concept for robust visual sign language recognition, which represents the recent developments in this field. The proposed recognition system aims for signer-independent operation and utilizes a single video camera for data acquisition to ensure user-friendliness. Since sign languages make use of manual and facial means of expression, both channels are employed for recognition.

Keywords:

Sign language detection, Sign language processing, Deaf-dumb hearing impaired, intelligent systems

# Table of Contents

|  |  |
| --- | --- |
| Abstract | i |
| 1.    Introduction |  |
| 1.1  Problem Definition | 1 |
| 1.2 Project Overview | 2 |
| 1.3 Hardware Specification | 3 |
| 1.4 Software Specification | 3 |
| 2.    Literature Survey | 4 |
| 2.1 Existing System | 6 |
| 2.2 Proposed System | 7 |
| 2.3 Literature Review Summary | 8 |
| 3.    Problem Formulation | 9 |
| 4.    Research Objective | 10 |
| 5.    Methodologies | 11 |
| 6.    Conclusion | 13 |
| 7.    Tentative Chapter Plan for the proposed work | 14 |
| 8.    Reference | 15 |
|  |  |

# 1. INTRODUCTION

## 1.1 Problem Definition

Deaf and hard-of-hearing persons, as well as others who are unable to communicate verbally, utilise sign language to communicate within their communities and with others. Sign languages are a set of present languages that communicate information using a visual-manual modality. The dilemma of real-time finger-spelling recognition in Sign Language is discussed. We gathered a dataset for identifying 36 distinct gestures (alphabets and numerals) and a dataset for typical hand gestures in ISL created from scratch using webcam images. The system accepts a hand gesture as input and displays the identified character on the monitor screen in real time. This project falls under the category of human-computer interaction (HCI) and tries to recognise multiple alphabets (a-z), digits (0-9) and several typical ISL hand gestures. To apply Transfer learning to the problem, we used a Pre-Trained SSD Mobile net V2 architecture trained on our own dataset. In the vast majority of situations, we constructed a robust model that consistently classifies Sign language.

This project falls within the HCI (Human Computer Interface) sector and seeks to recognise multiple alphabets (a-z), digits (0-9) and several typical ISL family hand motions such as Thank you, Hello, and so on. Hand-gesture recognition is a difficult problem, and ISL recognition is particularly difficult owing to the use of both hands. Many studies have been done in the past employing sensors (such as glove sensors) and various image processing techniques (such as edge detection, Hough Transform, and so on), but they are quite costly, and many people cannot afford them.

## 1.2 Problem Overview

Many people in India are deaf or hard of hearing, thus they communicate with others using hand gestures. However, aside from a small group of people, not everyone is familiar with sign language, and they may need an interpreter, which may be complex and costly. The goal of this research is to build software that can anticipate ISL alphanumeric hand movements in real time, bridging the communication gap.

Sign language is largely used by the disabled, and there are few others who understand it, such as relatives, activists, and teachers at SekolahLuarBiasa (SLB). Natural gestures and formal cues are the two types of sign language[1]. The natural cue is a manual (hand-handed) expression agreed upon by the user (conventionally), recognised to be limited in a particular group (esoteric), and a substitute for words used by a deaf person (as opposed to body language). A formal gesture is a cue that is established deliberately and has the same language structure as the community's spoken language.

Sign language is a visual language. It mainly consists of 3 major components:

1.Fingerspelling: Spell out words character by character, and word level association which involves hand gestures that convey the word meaning. The static Image Dataset is used for this purpose.

2.World-level sign vocabulary: The entire gesture of words or alphabets is recognized through video classification. (Dynamic Input / Video Classification)

3.Non-manual features: Facial expressions, tongue, mouth, body positions

## 1.3 Hardware Specification

1.Interface: Jupyter notebook for inserting python libraries in a notebook format, it is typically a python code where we can easily estimate our data sets model in one single notebook.

2.Operating System Environment: Windows 10

3.Hardware Environment: RAM- 16GB ,GRAPHIC CARD – 6GB , ROM-1060TB

## 1.4 Software Specification

Software: Python (3.7.4), Anaconda(2019-0.7) ,IDE (Jupyter), Numpy (version 1.16.5), cv2 (openCV) (version 3.4.2) , Tensorflow (version 2.0.0) , Github , Virtual Studio (2022) ,CUDA(10.1) and CuDNN(7.6) (For NIVIA GPU for faster training model) ,Protoc

# 2. LITERATURE SURVEY

Starner and Pentland[10] provided one of the early studies on sign language recognition. They demonstrated a real-time hidden Markov model-based system that detected sentence level American Sign Language (ASL) movements with the help of a webcam. They described two experiments: the first uses a desk-mounted camera to view the user, while the second uses a camera embedded in the user's cap.

The authors in paper [3] presented a system for the automatic translation of the gestures of the manual alphabets in Arabic Sign Language. This system made use of images of the gesture as input which were then processed and converted into a set of features that comprised of some length measures which indicated the fingertip’s position. The subtractive clustering algorithm and the least-squares estimator were used for classification. The system achieved an accuracy of 95.55%.

In [4] Nadia R. Albelwi and Yasser M. Alginahi proposed a real-time Arabic Sign Language system where a video camera was used to capture real-time video as an input to the system. The authors used a Haar-like algorithm to track the hand in the video frames and applied preprocessing techniques like skin detection and size normalization to extract the region of interest. To obtain the feature vectors, Fourier Transformation is applied to the resultant images which are transformed into the frequency domain. The classification is performed using the k-Nearest Neighbor (KNN) algorithm and the system achieves an accuracy of 90.55%.

In [5] Balakrishnan, G., P. S. Rajam, et al., proposed a system that converts a set of 32 combinations of the binary number, which represents the UP and DOWN positions of the five fingers into decimal. The binary numbers are first converted into a decimal form by using the binary-decimal conversion algorithm and then the decimal numbers are converted to their corresponding Tamil letters. Static images of the gesture were used as the input to the system where a canny-edge detection algorithm was applied to extract the edges of the palm and Euclidean Distance was applied to identify the position of the fingers. The system achieved an accuracy of 98.75%.

The authors in paper [6] proposed a system employing bspline approximation to develop a novel vision-based recognition system for Indian sign language alphabets and digits. By using the Maximum Curvature Points (MCPs) as Control points, their technique approximates the extracted boundary from the region of interest to a B-Spline curve. The B-spline curve is then smoothed iteratively, resulting in the extraction of Key Maximum Curvature Points (KMCPs), which are the major contributors to the gesture shape. As a result, the spatial coordinates of the KMCPs in the 8 Octant Regions of the 2D Space that are given for classification yield a translation and scale-invariant feature vector. The accuracy of numbers was 93.2 percent, and the accuracy of alphabets was 91.83 percent.

In paper [7], the authors proposed a system that can recognize and convert ISL gestures from a video feed into English voice and text. They did this by segmenting the shapes in the video stream using several image processing techniques such as edge detection, wavelet transform, and picture fusion. Ellipsoidal Fourier descriptors were used to extract shape features, while PCA was utilized to optimize and reduce the feature set. The fuzzy inference system was used to train the system, which resulted in a 91% accuracy.

In paper [8], the authors suggested a method for automatically recognizing Indian sign language gestures. The proposed method employs digital image processing techniques and uses the YCbCr color space for hand detection, with the input image being transformed beforehand. Distance transformation, Projection of distance transformation coefficients, Fourier descriptors, and feature vectors are some of the techniques used to extract the features. An artificial neural network was used to classify the data, and the recognition rate was 91.11 percent.

## 2.1 Existing System

In existing system the module was developed for dumb person using flex sensor, there user hand is attached with the flex sensors. On this module the flex sensor reacts on bend of each finger individually. By taking that value controller starts to react with speech, each flex sensor holds unique voice stored in APR Kit and for each sign it will play unique voice. And in other existing system, the work is done only for some alphabets and not for the words or sentences, and accuracy obtained is very low.

**Limitations of existing system**

* In existing system it’s restricted to only 10 voice announcements it may reduce product capacity
* One of the major problems of the existing system is Dumb person should always carry the hardware with him
* User can’t do any other work with flex sensor on fingers and also sensors should be placed straight
* The controller may think that the user is giving command and finally it may result in unwanted results and less hardware lifetime

## 2.2 Proposed System

In the proposed system the unable or dumb person should provide a gesture or sign image to the system. The system evaluates the sign input with matlab image processing technique and classifies the input to the recognized identification. Later it initiates the voice media through the system when the input image matches with the given dataset. And the output will be shown in the text format too. This is a prototype to develop the concept of converting the sign language to speech and text. The aim of this paper is to provide an application to the society to establish the ease of communication between the deaf and mute people by making use of image processing algorithm.

**Advantages of proposed system**

* When comparing with existing system user can give more signs
* The module provides two way communications which helps in easy interaction between the normal people and disables
* Easy to Interface
* Flexible

## 2.3 Literature Review Summary

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and**  **Citation** | **Article/ Author** | **Tools/ Software** | **Technique** | **Source** | **Evaluation Parameter** |
| 2022 | Real Time Sign Language Recognition System for Hearing and Speech Impaired People | Python Machine Learning | CNN | http://surl.li/fhxbe | Accuracy of 80% |
| 2022 | Deepsign: Sign Language Detection and Recognition | Python Deep Learning | feedback-based learning models | https://rb.gy/ifvz7j | 87% Accuracy |
| 2019 | Sign Language Detection “in the Wild” | Python Machine Learning | RNN | https://rb.gy/tqutsj | Precision of 83% |
| 2021 | ML Based Sign Language Recognition System | Python Machine Learning | KNN | https://rb.gy/qejolc | 65% Accuracy. |
| 2017 | Machine Learning Techniques for Indian Sign Language Recognition | Python Machine Learning | ML Algorithms | https://rb.gy/gakute | 90% |

# 3. PROBLEM FORMULATION

Conversing with people having a hearing disability is a major challenge. Deaf and Mute people use hand gesture sign language to communicate, hence normal people face problems in recognizing their language by signs made. Hence there is a need for systems that recognize the different signs and conveys the information to normal people.

The solution is to develop a translator that can detect sign language used by a disabled person, and then feed that sign into a machine-learning algorithm called transfer learning, which is then detected by the neural network and translated on the screen so that a normal person can understand what the sign is saying.

It's a lot easier now, thanks to speech to text and translators. But what about individuals who are unable to speak or hear?   The main goal of this project is to create an application that can assist persons being unable to speak or hear. The language barrier is also a very significant issue. Hand signals and gestures are used by people who are unable to speak. Ordinary people have trouble comprehending their own language. As a result, a system that identifies various signals and gestures and relays information to ordinary people is required. It connects persons who are physically handicapped with others who are not.

Many firms are creating solutions for deaf and hard of hearing persons, but not everyone can afford them. Some are very pricey for ordinary middle-class individuals to bring.

# 4. OBJECTIVES

The proposed work is aimed to carry out work leading to the development of an approach for SIGN LANGUAGE DETECTION MODEL. More than 360 million of world population suffers from hearing and speech impairments [3]. Sign language detection is a project implementation for designing a model in which web camera is used for capturing images of hand gestures which is done by open cv.

After capturing images, labelling of images are required and then pre trained model SSD Mobile net v2 is used for sign recognition. Thus, an effective path of communication can be developed between deaf and normal audience. Three steps must be completed in real time to solve our problem:

1. Obtaining footage of the user signing is step one (input).

2. Classifying each frame in the video to a sign.

3. Reconstructing and displaying the most likely Sign from classification scores (output).

# 5. METHODOLOGY

The following methodology will be followed to achieve the objectives defined for proposed research work:

Phase1: Searching research papers and collecting data of Sign Language Detection, software requirements.

Phase2: Implementation of code of Sign Language Detection and Data Gathering and Train the Recognizer

Phase3: Implementation of code of Model

Phase4: Finalize Project and All Documentation of project.

* Fundamental steps in image processing are:

1. Image acquisition: to acquire a digital image

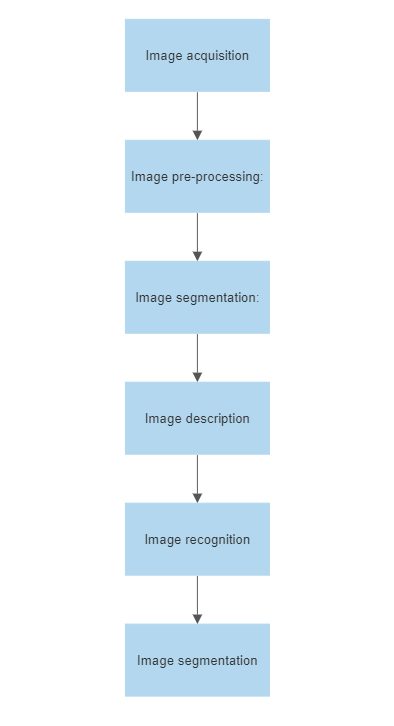
2. Image pre-processing: to improve the image in ways that increases the chances for success of the other processes.

3. Image segmentation: to partitions an input image into its constituent parts of objects.

4. Image description: to extract the features that result in some quantitative information of interest of features that are basic for differentiating one class of objects from another.

5. Image recognition: to assign a label to an object based on the information provided by its description.

6. Image segmentation: to convert the input data to a from suitable for computer processing.



# 6.CONCLUSION

Sign languages are kinds of visual languages that employ movements of hands, body, and facial expression as a means of communication. Sign languages are important for specially-abled people to have a means of communication. Through it, they can communicate and express and share their feelings with others. The drawback is that not everyone possesses the knowledge of sign languages which limits communication. This limitation can be overcome by the use of automated Sign Language Recognition systems which will be able to easily translate the sign language gestures into commonly spoken language. In this paper, it has been done by TensorFlow object detection API. The system has been trained on the Indian Sign Language alphabet dataset. The system detects sign language in real-time. For data acquisition, images have been captured by a webcam using Python and OpenCV which makes the cost cheaper. The developed system is showing an average confidence rate of 85.45%. Though the system has achieved a high average confidence rate, the dataset it has been trained on is small in size and limited. In the future, the dataset can be enlarged so that the system can recognize more gestures. The TensorFlow model that has been used can be interchanged with another model as well. The system can be implemented for different sign languages by changing the dataset.

## 7. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

**CHAPTER 1: INTRODUCTION**

This chapter will cover the overview of the existing models and the need of better model to detect sign language.

**CHAPTER 2: LITERATURE REVIEW**

This chapter include the literature available for SIGN LANGUAGE DETECTION MODEL. The findings of the researchers will be highlighted which will become basis of current implementation.

**CHAPTER 3: OBJECTIVE**

This chapter will provide introduction to the concepts which are necessary to understand the proposed system.

**CHAPTER 4: METHODOLOGIES**

This chapter will cover the technical details of the proposed approach.

**CHAPTER 5: EXPERIMENTAL SETUP**

This chapter will provide information about the subject system and tools used for evaluation of proposed method.

**CHAPTER 6: CONCLUSION AND FUTURE SCOPE**

The result of proposed technique will be discussed in this chapter.

## REFERENCES

[1] Noor Adnan Ibraheem, RafiqulZaman Khan, 2012, Survey on Various Gesture Recognition Technologies and Techniques, International Journal of Computer Applications, Volume 50, No. 7

[2] G. R. S. Murthy & R. S. Jadon, 2009. A Review of Vision Based Hand Gestures Recognition, International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405-410.

[3] Al-Jarrah, A. Halawani, “Recognition of gestures in Arabic sign language using neuro-fuzzy systems,” The Journal of Artificial Intelligence 133 (2001) 117–138..

[4] Nadia R. Albelwi, Yasser M. Alginahi, “Real-Time Arabic Sign Language (ArSL) Recognition“ International Conference on Communications and Information Technology 2012.

[5] Rajam, Subha & Ganesan, Balakrishnan. (2011). Real time Indian Sign Language Recognition System to aid deaf-dumb people. International Conference on Communication Technology Proceedings, ICCT. 10.1109/ICCT.2011.6157974.

[6] Geetha M., Manjusha U. C. 2013,A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation, International Journal of Computer Science and Engineering (IJCSE)

[7] P.V.V Kishore, P. Rajesh Kumar, E. Kiran Kumar & S.R.C.Kishore, 2011, Video Audio Interface for Recognizing Gestures of Indian Sign Language, International Journal of Image Processing (IJIP), Volume 5, Issue 4, 2011 pp. 479-503

[8] Adithya, V., Vinod, P. R., Gopalakrishnan, U., and Ieee (2013). "Artificial Neural Network Based Method for Indian Sign Language Recognition."

[9] http://timesofindia.indiatimes.com/home/education/n ews/realising-the-importance-of-indian-sign-languagedictionary/articleshow/86221166.cms

[10] T. Starner and A. Pentland, "Real-time American sign language recognition from video using hidden markov models", Technical Report, M.I.T Media Laboratory Perceptual Computing Section, Technical Report No. 375, 199