ABSTRACT

Sign language is an overlooked concept even though there being a large social group which could benefit by it. Not everyone knows how to interpret a sign language when having a conversation with a deaf and dumb person. There is always a need to communicate using sign language. One finds it hard to communicate without a translator. To solve this, we need a common translator that is understood by common people and will help them to communicate without any barriers. Image classification and machine learning can be used to help computers recognize sign language, which could then be interpreted by other people. Pre-processing will be performed on images to get cleaned input. After that convolutional neural network (CNN) will be used to recognize sign language gestures. The main aim of this project is to eliminate the barrier between the deaf and dumb and the rest.

Table of Contents

1. INTRODUCTION

1.1 OUTLINE OF THE PROJECT

Indian sign language is a predominant sign language Since the only disability D&M people have is communication related and they cannot use spoken languages hence the only way for them to communicate is through sign language. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behaviour and visuals. Deaf and dumb(D&M) people make use of their hands to express different gestures to express their ideas with other people. Gestures are the nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language.

American sign language is a predominant sign language Since the only disability D&M people have is communication related and they cannot use spoken languages hence the only way for them to communicate is through sign language. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behaviour and visuals. Deaf and dumb(D&M) people make use of their hands to express different gestures to express their ideas with other people. Gestures are the nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language.

For interaction between normal people and D&M people a language barrier is created as sign language structure which is different from normal text. So, they depend on vision-based communication for interaction. If there is a common interface that converts the sign language to text, the gestures can be easily understood by the other people. So, research has been made for a vision-based interface system where D&M people can enjoy communication without really knowing each other's language. The aim is to develop a user-friendly human computer interface (HCI) where the computer understands the human sign language. There are various sign languages all over the world, namely American Sign Language (ASL), French Sign Language, British Sign Language (BSL), Indian Sign language, Japanese Sign Language and work has been done on other languages all around the world

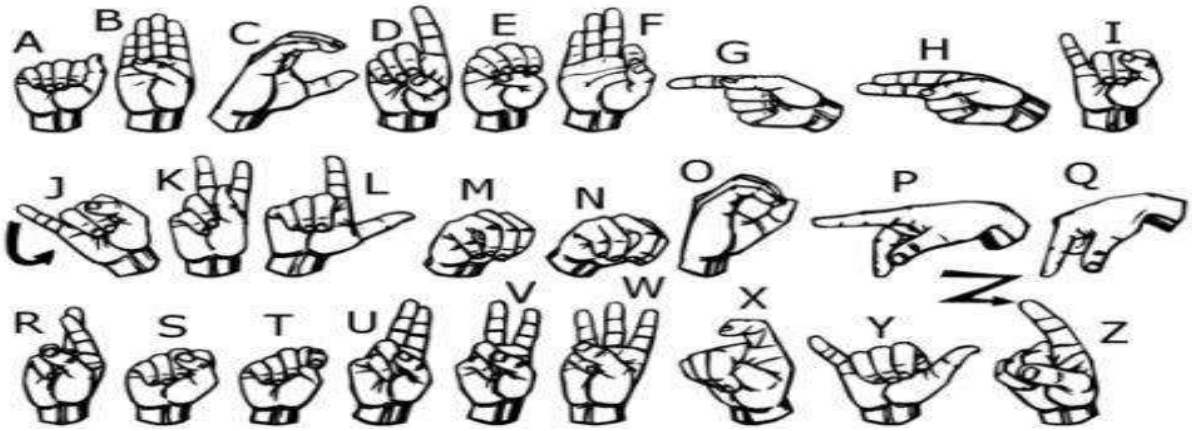
The importance of the very use of this method is increasing day by day as this gives the people with disabilities and opportunity to try their hands in different fields that require communication. With the help of this project, they can communicate with the majority in any industry, thus giving them an even playing field to cope up with. Here, we use the method of object identification followed by recognition to help us differentiate between the many symbols used in the sign.

1.2 LITERATURE VIEW

Sign language is a visual language and consists of 3 major components

Fingerspelling	Word level sign vocabulary	Non-manual features
Used to spell words	Used for the majority of communication	Facial expressions and tongue, mouth, and body position.

In our project we basically focus on producing a model which can recognise Fingerspelling based hand



gestures in order to form a complete word by combining each gesture. The gestures we aim to train are as given in the image below.

In the recent years there has been tremendous research done on the hand gesture recognition.

With the help of literature survey done we realized the basic steps in hand gesture recognition are:-

i. Data-acquisition
ii. Data-pre-processing
iii. Gesture-classification

Data acquisition:

The different approaches to acquire data about the hand gesture can be done in the following ways:

1.Use of sensory devices

It uses electromechanical devices to provide exact hand configuration, and position. Different glove-based approaches can be used to extract information. But it is expensive and not user friendly.

2.Vision based approach

In vision-based methods computer camera is the input device for observing the information of hands or fingers. The Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices. These systems tend to complement biological vision by describe in artificial vision systems that are implemented in software and/or hardware.

The main challenge of vision-based hand detection is to cope with the large variability of human hand's appearance due to a huge number of hand movements, to different skin colour possibilities as well as to the variations in viewpoints, scales, and speed of the camera capturing the scene.

Data pre-processing and Feature extraction for vision-based approach:

- In [1] the approach for hand detection combines threshold-based colour detection with background subtraction. We can use Ad boost face detector to differentiate between faces and hands as both involve similar skin-color.

- We can also extract necessary image which is to be trained by applying a filter called Gaussian blur. The filter can be easily applied using open computer vision also known as OpenCV and is described in[3].

- For extracting necessary image which is to be trained we can use instrumented gloves as mentioned in [4]. This helps reduce computation time for pre-processing and can give us more concise and accurate data compared to applying filters on data received from video extraction.

- We tried doing the hand segmentation of an image using colour segmentation techniques but as mentioned in the research paper skin colour and tone is highly dependent on the lighting conditions due to which output, we got for the segmentation we tried to do were no so great. Moreover we have a huge number of symbols to be trained for our project many of which look similar to each other like the gesture for symbol 'V' and digit '2', hence we decided that in order to produce better accuracies for our large number of symbols, rather than segmenting the hand out of a random background we keep background of hand a stable single colour so that we don't need to segment it on the basis of skin colour. This would help us to get better results.

Gesture Classification:

- In [1] Hidden Markov Models (HMM) is used for the classification of the gestures. This model deals with dynamic aspects of gestures. Gestures are extracted from a sequence of video images by tracking the skin-color blobs corresponding to the hand into a body– face space cantered on the face of the user. The goal is to

recognize two classes of gestures: deictic and symbolic. An image is filtered using a fast look–up indexing table. After filtering, skin colour pixels are gathered into blobs. Blobs are statistical objects based on the location (x, y) and the colorimetry (Y,U,V) of the skin colour pixels in order to determine homogeneous areas.

- In Naïve Bayes Classifier is used which is an effective and fast method for static hand gesture recognition. It is based on classifying the different gestures according to geometric based invariants which are obtained from image data after segmentation. Thus, unlike many other recognition methods, this method is not dependent on skin colour. The gestures are extracted from each frame of the video, with a static background. The first step is to segment and label the objects of interest and to extract geometric invariants from them. Next step is the classification of gestures by using a K nearest neighbour algorithm aided with distance weighting algorithm (KNNDW) to provide suitable data for a locally weighted Naïve Bayes “classifier.

- According to paper on “Human Hand Gesture Recognition Using a Convolution Neural Network” by Hsien-I Lin , Ming-Hsiang Hsu, and Wei-Kai Chen graduates of Institute of Automation Technology National Taipei University of Technology Taipei, Taiwan, they construct asking model to extract the handout of an image and then apply binary threshold to the whole image. After obtaining the threshold image they calibrate it about the principal axis in order to centre the image about it. They input this image to a convolutional neural network model in order to train and predict the outputs. They have trained their model over 7 hand gestures and using their model they produce an accuracy of around 95% for those 7 gestures.

1.3 SYSTEM IMPLEMENTATION

For interaction between normal people and D&M people a language barrier is created as sign language structure which is different from normal text. So, they depend on vision-based communication for interaction.

If there is a common interface that converts the sign language to text the gestures can be easily understood by the other people. So, research has been made for a vision-based interface system where D&M people can enjoy communication without really knowing each other's language.

If there is a common interface that converts the sign language to text the gestures can be easily understood by the other people. So, research has been made for a vision-based interface system where D&M people can enjoy communication without really knowing each other's language.

1.4 OBJECTIVES

1) Decision making power is provided by this system.
2) Accurate result can be obtained.

3) This system makes selection process more effective.

4) To increase efficiency proposed system is dependent on classification method.

5) Proposed system is used to reduce confusion at the time of processing data average.

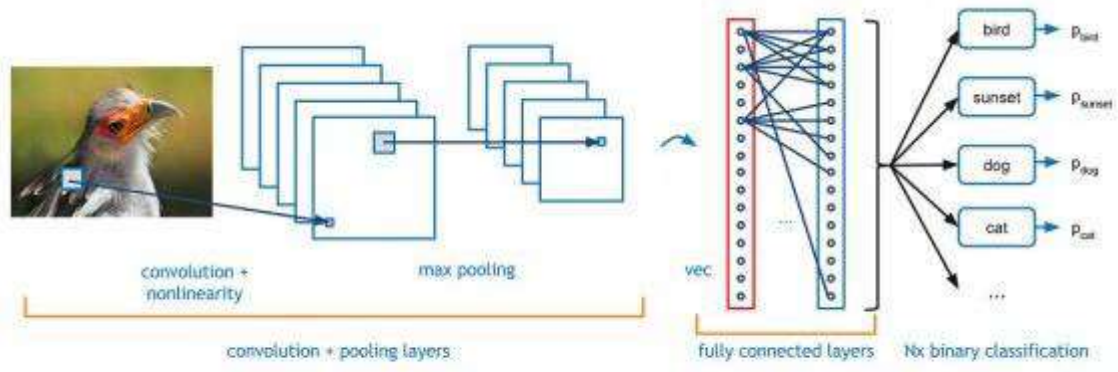


Figure 5.2: Convolution neural networks

2. AIM AND SCOPE OF THE PRESENT INVESTIGATION

2.1 SIGN LANGUAGE RECOGINITION

Sign language recognition and translation is a research area with high potential impact. There are over 300 sign languages used around the world, and 70 million deaf people are using them. Sign language processing would break down all the barriers for sign language users. This can be very helpful for the deaf and dumb people in communicating with others as knowing sign language is not something that is common to all, moreover, this can be extended to creating automatic editors, where the person can easily write by just their hand gestures. The types of data available and the relative merits are explored allowing examination of the features which can be extracted. Classifying the manual aspects of sign (similar to gestures) is then discussed from a tracking and non-tracking viewpoint before summarizing some of the approaches to the non-manual aspects of sign languages. Methods for combining the sign classification results into full SLR are given showing the progression towards speech recognition techniques and the further adaptations required for the sign specific case. Finally, the current frontiers are discussed, and the recent research presented. This covers the task of continuous sign recognition, the work towards true signer independence, how to effectively combine the different modalities of sign, making use of the current linguistic research and adapting to larger more noisy data sets.

2.2 SERVERPORTAL

Gesture recognition information system is a project of the Department of Intelligent Systems in Control and Automation, which includes the development of subsystem for recognizing various types of communication between people with hearing and voice disabilities, fingerprints, gestures using the user's hands and his emotions. The recognition system is a combination of software and hardware tools that allow remotely using the recognition system.

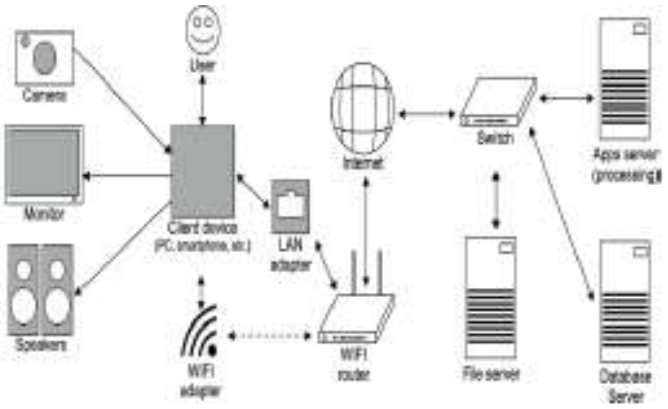


Figure:2

3. SYSTEM REQUIRMENT

3.1 REQUIRMENTS SPECIFICATION

3.1.1 HARDWARE REQUIRMENTS

Processor – Intel i5

Ram – 4GB

Hard Disk Dive – 40GB

Monitor – LCD

3.1.2 SOFTWARE REQUIRMENTS

Operating System – Windows 10
Software – SPYDER

3.2 ABOUT THE SOFTWARE

3.2.1 SPYDER

Spyder, the Scientific Python Development Environment, is a free integrated development environment (IDE) that is included with Anaconda. It includes editing, interactive testing, debugging, and introspection features.

Spyder is an open-source cross-platform IDE. The PYTHON Spyder IDE is written completely in Python. It is designed by scientists and is exclusively for scientists, data-analyst, and engineers. It is also known as the Scientific Python Development IDE and has a huge set of remarkable features,

WHY SPYDER?

i) It's an open-source cross-platform IDE for data science.
ii) It integrates the essentials libraries for data science, such as NumPy, SciPy, Matplotlib and I-Python, besides that, it can be extended with plugins.

iii) Spyder contains features like a text editor with syntax highlighting, code completion and variable exploring, which you can edit its values using a Graphical User Interface (GUI).

4. EXPERIMENTAL OR MATERIALS AND METHODS ALGORITHM USED

4.1 METHODOLOGY

Data Pre-processing Using Label-Img.
Object Detection using TensorFlow.
Real Time Detection.

A. DATA ACQUISITION

The different approaches to acquire data about the hand gesture can be done in the following ways:

1.USE OF SENSORY DEVICES

It uses electromechanical devices to provide exact hand configuration, and position. Different glove-based approaches can be used to extract information. But it is expensive and not user friendly.

2.VISION BASED APPROACH

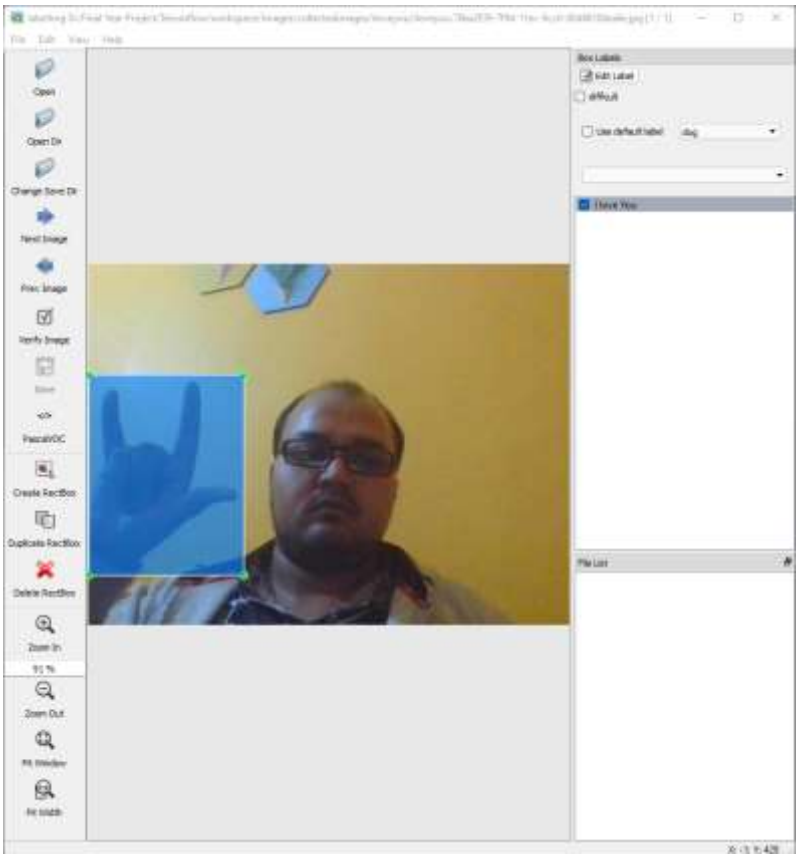
In vision-based methods computer camera is the input device for observing the information of hands or fingers. The Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices. These systems tend to complement biological vision by describing 7 artificial vision systems that are implemented in software and/or hardware. The main challenge of vision-based hand detection is to cope with the large variability of human hand's appearance due to a huge number of hand movements, to different skin-color possibilities as well as to the variations in viewpoints, scales, and speed of the camera capturing the scene.

However, due to the rarity of the chosen project, there were a little to no datasets available online. As a result, we had to make our own datasets to help train the code to give us the best results possible. It was made possible with the use of Open computer vision also known as the Open CV library to obtain the customized datasets we use in this program.

Fig 4.1. Image captured using figure.

B. DATA PREPROCESSING USING LABELING

After you train the algorithm to help us identify sign language, we need to make sure the algorithm is able to ignore the backgrounds and focus solely on the given dataset. Labelling is a graphical image annotation tool which is used for bounding boxes in images.



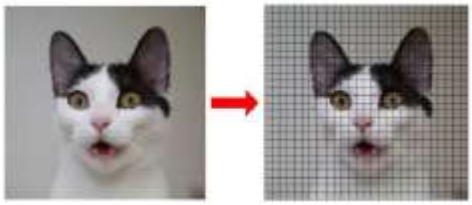
C. OBJECT DETECTION

Object detection is a method in which a program or a software can trace and detect the object from a given photo or any other visual data. The special attribute about object detection is that it identifies the class of object (person, table, chair, etc.) and their location-specific coordinates in the given image.

A bounding box around the object is used to indicate its location. The drawn bounding box may or may not pinpoint the exact location of the given piece. The potential to spot the provided object inside a photo explain the performance of the algorithm that has been used for detection. Sign language detection is one of the examples of object detection.

Object detection steps: -

Generation of small portion in the input as shown in the image below. As shown below in the figure large set of boxes are spanning the full image.



Feature extraction is the next step that will be done for each rectangular area to check whether or whether not the rectangle area contains a valid object.