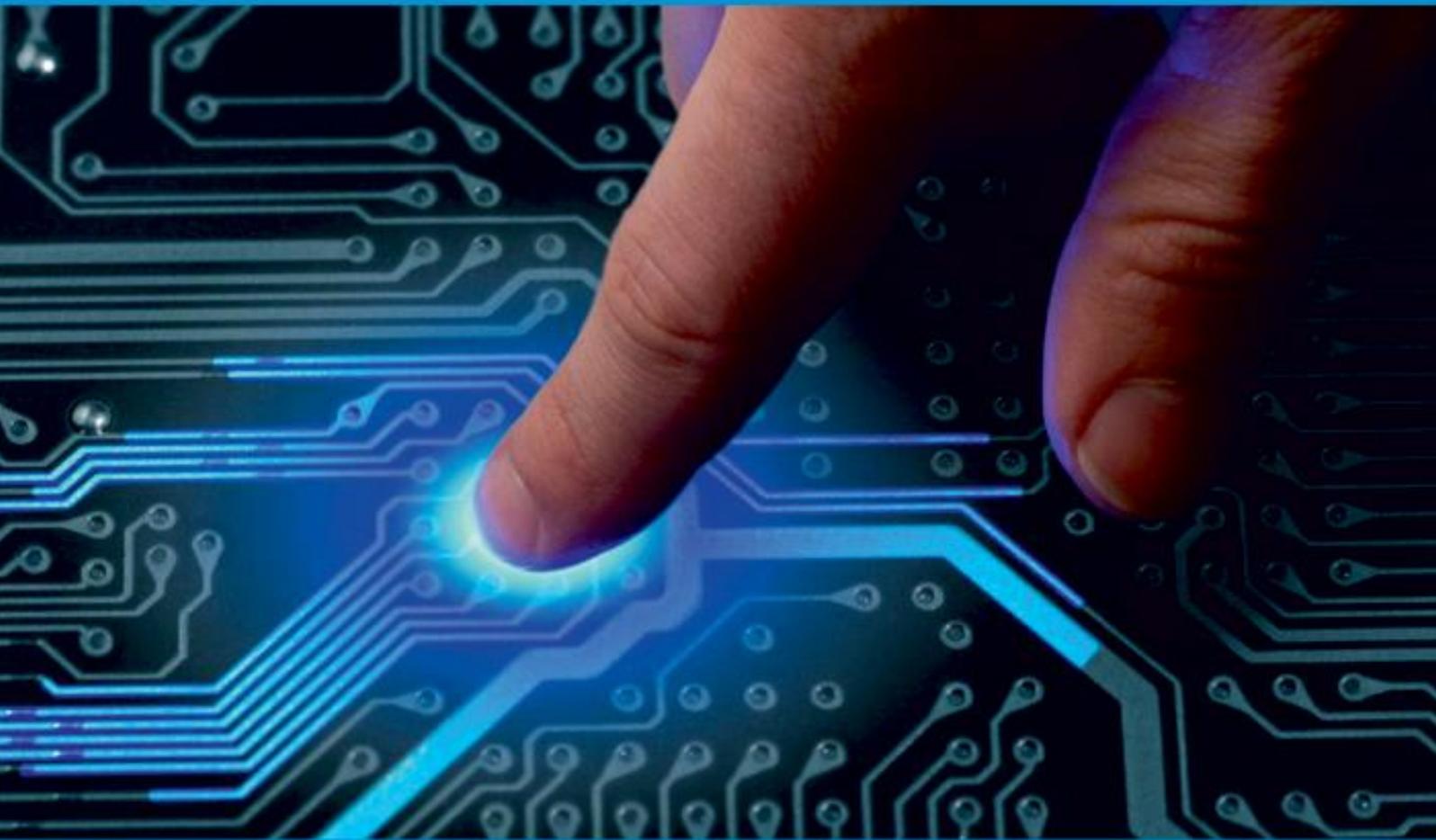




**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.379**

 9940 572 462

 6381 907 438

 [ijircce@gmail.com](mailto:ijircce@gmail.com)

 [www.ijircce.com](http://www.ijircce.com)

# S2S Translator Using CNN

Mrs. A. Subha M.E, M. Tenmoli, M. Priya dharshini, Y. Mounith, R. Harsha Vardhan

Associate Professor, Dept. of CSE., JNTUA University, Kuppam Engineering college, Andhra Pradesh, India

UG Students, Dept. of CSE., JNTUA University, Kuppam Engineering college, Andhra Pradesh, India

**ABSTRACT:** Sign language is a language used by dumb and deaf people to communicate with normal people. Normal people use sounds, unlike them, this language uses visual communication to convey the thoughts of dumb and deaf people. Sign language is achieved by continuously showing hands, the orientation of fingers, and facial expressions. In this project, we will develop a programmatic model that converts voice to sign language and also sign language to voice/text. We may be using different APIs (Python modules or Google API) and natural language processing semantics to break the text into a large number of smaller understandable words which require machine learning as a part. Predefined alphabet signs are given as inputs to the model. So this can use Artificial Intelligence technology audio into sign language and sign language to text.

**PROJECT OVERVIEW:**In this project, we will try to create an accurate model which would classify the sign language inputs to voice output and also voice input to sign language output in whichever way necessary. We will be using the concrete concepts of machine learning which are image processing and speech recognition according to our use case. Data sets of predefined sign language are used as the input to display the converted audio using speech recognition into the the sign language. Similarly image processing will be used to convert the sign language to speech/text.

## I. INTRODUCTION

In today's diverse and inclusive world, communication barriers can pose significant challenges for individuals with disabilities, particularly those who are deaf or hard of hearing. Speech-to-sign and vice versa translation systems play a vital role in bridging these communication gaps by enabling seamless interaction between individuals who use spoken language and those who use sign language. This introduction delivers into the importance, challenges, and feasibility of developing such systems using ConvolutionalNeuralNetworks(CNN).

In an increasingly interconnected world, effective communication is not just a convenience but a necessity. However, for individuals who are deaf or hard of hearing, traditional modes of communication, primarily reliant on spoken language, can present significant barriers. Sign language, with its visual-gestural modality, serves as the primary means of communication for many in the deaf community. Bridging the gap between spoken language and sign language is crucial for fostering inclusivity and ensuring equal access to communication opportunities. This introduction explores the significance, challenges, and feasibility of developing speech-to-sign and vice versa translation systems using Convolutional Neural Networks (CNNs).

## II. SIGN LANGUAGE AND SPEECH

Sign language is a language used by dumb and deaf people for communication. Normal people use sounds, unlike them, this language uses visual communication to convey the thoughts of dumb and deaf people. Sign language is achieved by continuously showing hands, the orientation of fingers, and facial expressions.Each country has its own sign language. Britishers have British Sign Language (BSL) and Americans have ASL(American Sign Language). Both these languages are different and American Sign Language using people could not understand British Sign Language and vice-versa. Some countries adopt options of sign language in their sign languages. No one invented sign language. No one knows the exact birth of sign language. Some sources suggest sign Language was developed nearly 200 years ago by combining local sign languages and LSF(French Sign Language). Today's sign language includes some elements of French Sign Language and the original local sign languages. Over time, today's sign language evolved as a mature language.

Sign Language is a separate language and distinct from the English language. They still contain some similar signs, they'll now not be understood by every other's user. Sign Language may be a language fully separate and distinct from English. Sign language has all the options of the language, like pronunciation, word formation, and ordination. Every

language has different ways of signaling different functions, such as asking a help rather than making an order, languages differ in these cases. English speakers raise their voices while asking a question and adjusting order to the words. Sign Language using people to ask questions with their eyebrows by raising them, widening eyes, and tilting bodies forward direction. In different languages, specific ways in which of expressing ideas in sign communication vary. In addition to individual variations in expression, signing has regional accents and dialects. Just as certain English words are spoken differently in different parts of the country, sign language has regional variations in the rhythm of signing, pronunciation, slang, and signs used. Other sociological factors, including age and gender, can affect sign language usage and contribute to its variety, just as with spoken languages. Sign language contains figure-spelling and English letters are shown using fingers. In the finger-spelled alphabet, each letter has distinct hand shape and figure orientation.

### III. TECHNOLOGY

#### IMAGE PROCESSING

Image Processing means processing of a digital image by means of a computer. It can also be defined as a way of using computer algorithms, in order to get enhanced image either to extract some useful information or for some other kind of task.

Image processing mainly the following steps:

- └ 1. Importing the image
- └ 2. Analyzing manipulating the image
- └ 3. Output in which result can be an altered image of a report on the analyzed image.

Phases of image processing

**1. Acquisition:** The image acquisition process consists of three steps; energy reflected from the object of interest, an optical system which focuses the energy and finally a sensor which measures the amount of energy.

This is a simple task in which the following happens when an image is given as input:

- a) Scaling
- b) Colour conversion (RGB to Gary or Vice-Versa)

The operations carried out in those two can differ based on the use-case.

**2. Image enhancement:** In this step, we will be extracting some of the hidden details from an image and its subjective properties. Image enhancement is the procedure of improving the quality and information content of original data before processing. Common practices include contrast enhancement, spatial filtering, density slicing, and FCC. Contrast enhancement or stretching is performed by linear transformation expanding the original range of Gary level.

#### SPEECH RECOGNITION

Speech recognition is a subfield of computer science and linguistics that develops methodologies and technologies that enable the recognition and translation of spoken language into text by computers. It is also known as Automatic speech recognition (ASR), computer speech recognition or speech to text (STT).

#### Performance

The performance of any speech recognition systems is usually evaluated in terms of accuracy and speed. Accuracy is usually rated with word error rate (WER), whereas speed is measured with the real time factor.

Speech recognition by machine is a very complex problem, however vocalizations vary in terms of accent, pronunciation, articulation, roughness, nasality, pitch, volume and speed. Speech is distorted by a background noise and echoes, electrical characteristics. Accuracy of speech recognition may vary with the following:

- └ 1. Vocabulary size and confusability
- └ 2. Speaker dependence versus independence
- └ 3. Isolated, discontinuous or continuous speech
- └ 4. Task and language constraints
- └ 5. Read versus spontaneous speech
- └ 6. Adverse condition

### NEURAL NETWORKS

A neural network is composed nodes. A neural network is an artificial neural network, for solving Artificial Intelligence problems weights. .

CNN or convolutional neural network or ConvNet is a deep learning algorithm that takes an image as input and assigns importance to those images, so it can differentiate one from the other. The pre-processing in a CNN is very lower as compared with other classification algorithm.

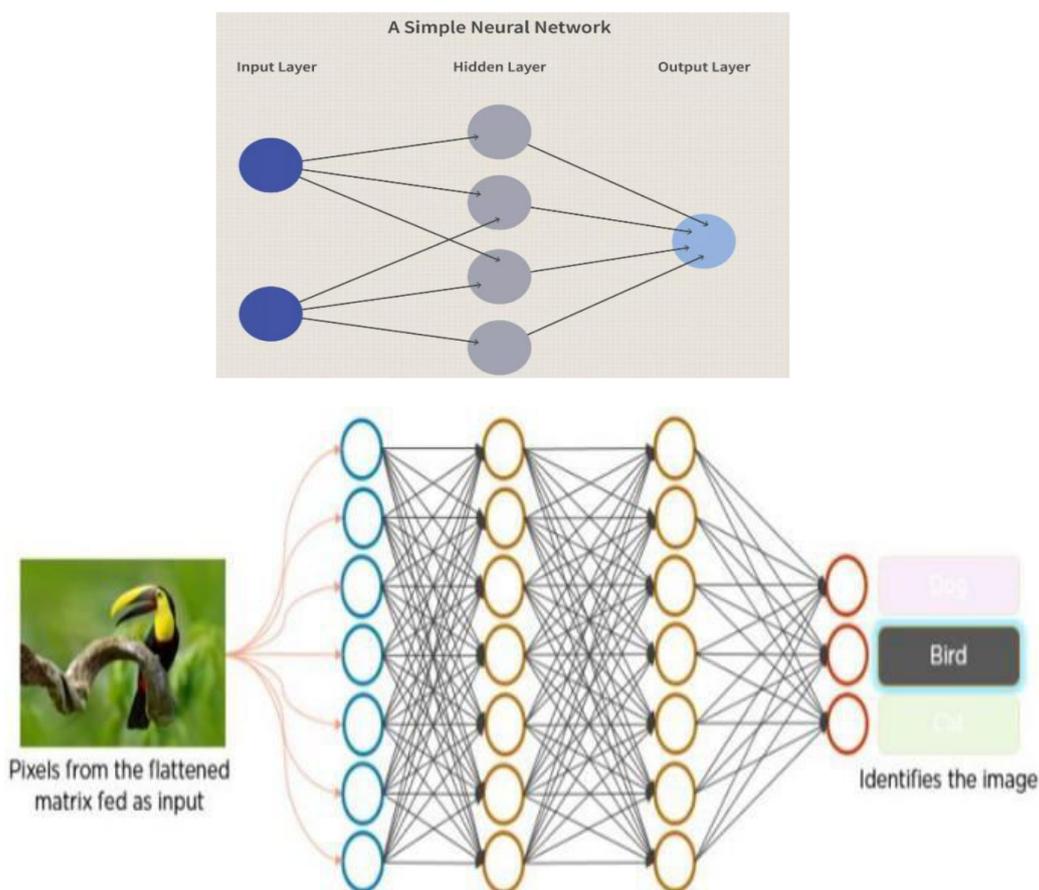


FIGURE 1. CNN reading the image of bird

The below figure shows the fingers of human hand.

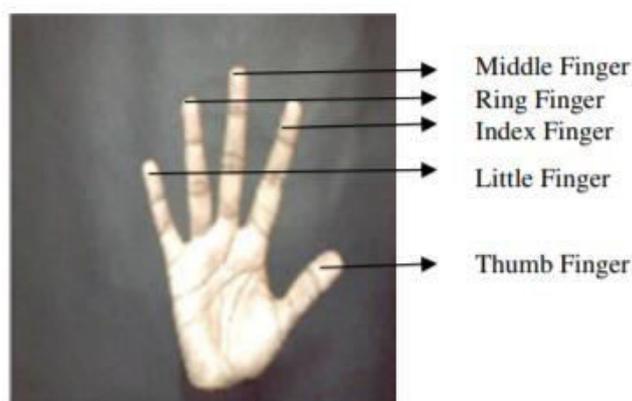


FIGURE 2. Name of fingers in hand

The below figure shows how the system process the image of the hand when you place in front of the camera. The difference between the original image, segmented image, and Edge detected image is clearly shown below.

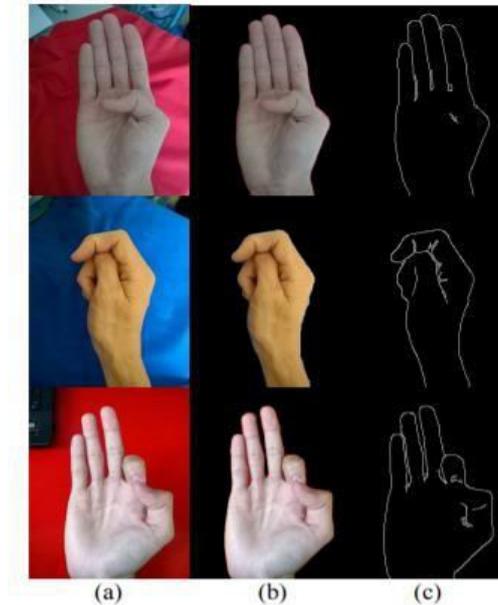


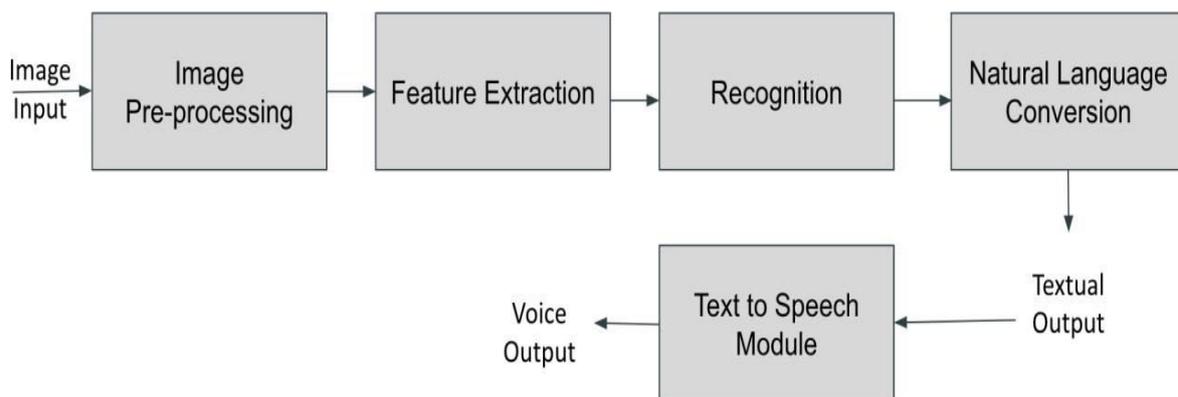
FIGURE 3. sign languages (a) original Image (b) segmented Image (c) Edge detected Image

#### IV.SYSTEM DESIGN

##### SYSTEM ARCHITECTURE

###### Image pre-processing:

Image processing can be defined as the processing of a digital image by with of a computer. It uses computer algorithms to get better image either to extract some useful information or for some other kind of task.



###### 1. Feature Extraction:

Feature extraction can be defined as reducing the dimensionality of an image in which the raw data is divided and reduced to more manageable groups.

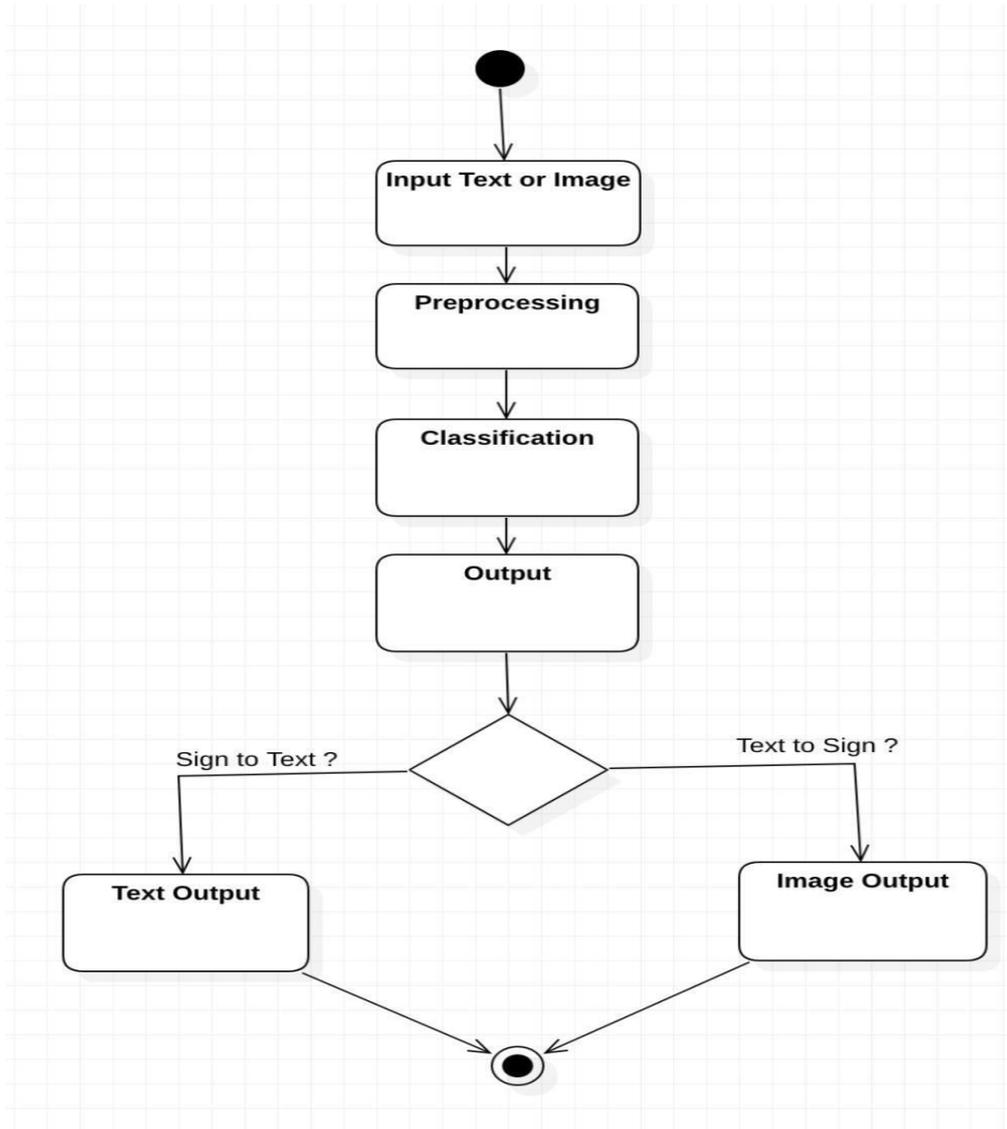
###### 2. Recognition:

Machine-based visual tasks are performed by image recognition. Labelling is given to the images with meta-tags, to performing image content search.

###### 3. Natural language conversion:

In natural language conversion, the recognized image is converted into a English text and text output is given as input to the text-to-speech module.

**ACTIVITY DIAGRAM**



**V. TESTING & OUTPUTS**

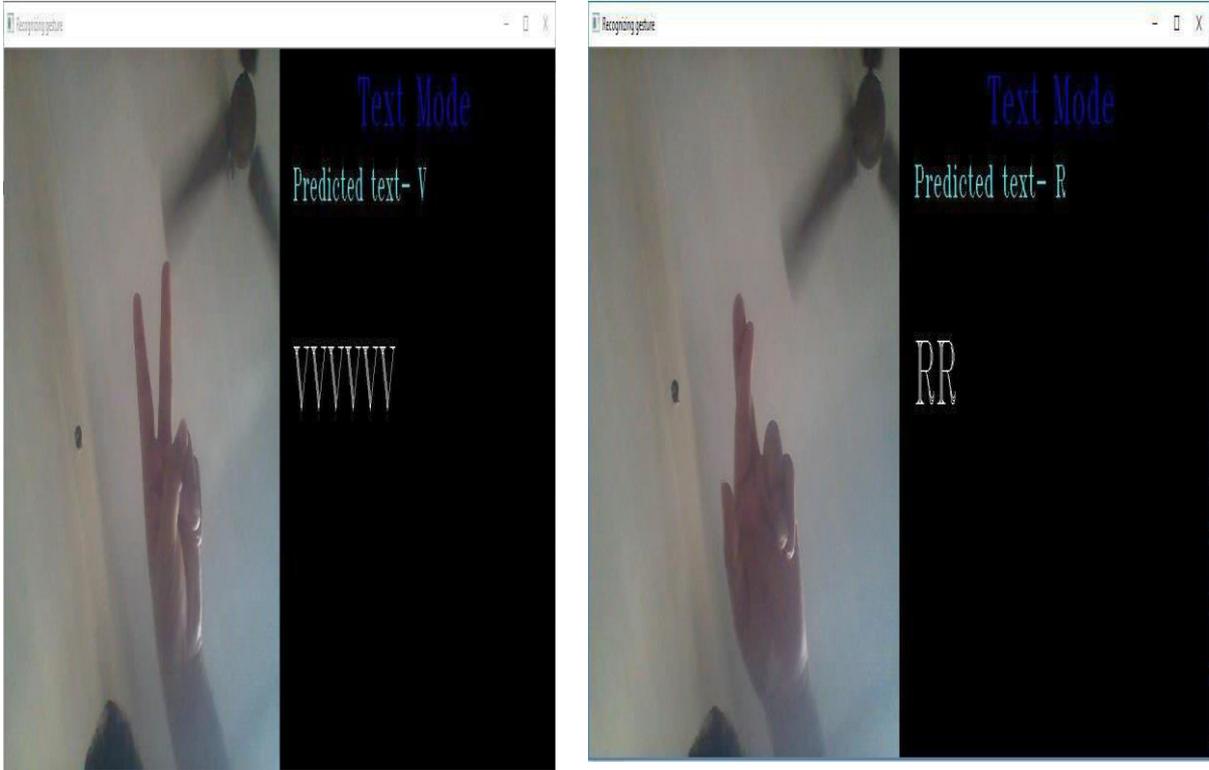
Testing is the process where the test data will get prepared and is employed for testing the modules individually and later the validation is given for the fields. Then the system testing takes place which makes sure that each and every component of the system property functions as a unit. The test data should be chosen in such a way that it should pass through all the possible circumstances. The following is the description of the total testing strategies, which were disbursed during the time of testing.

**SYSTEM TESTING**

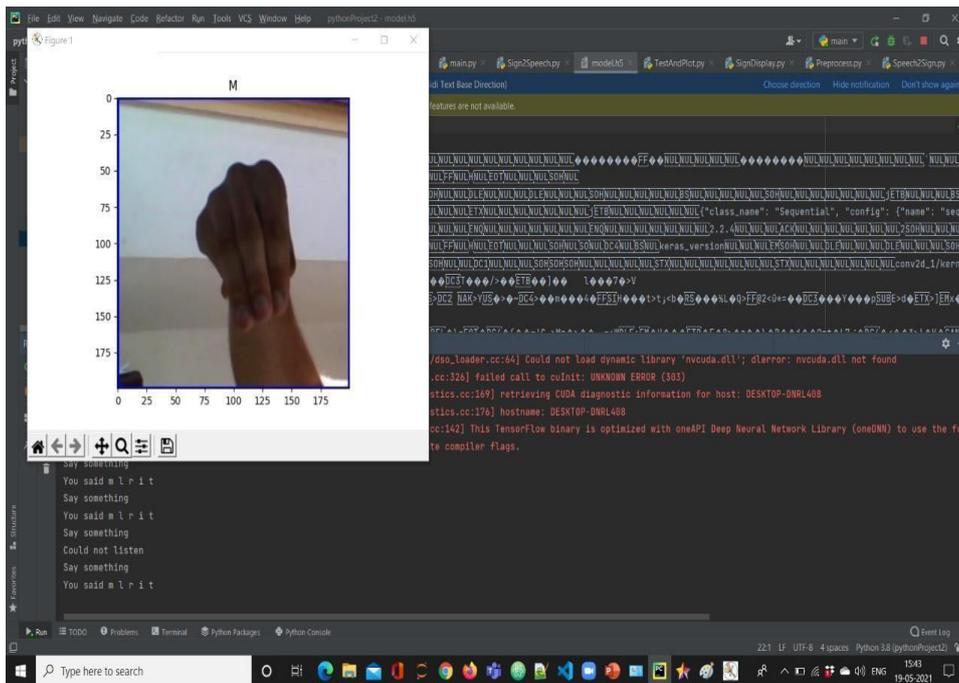
Testing has come an integral part of any project or system especially with in the field of an information technology. The importance of testing may be a method of justifying, if one is prepared to maneuverer further, be it to be check if one is capable to with stand the trials of a specific situation can't be underplayed and which is why testing before development is so critical. When the software is developed before it is given to user to user the software must be tested whether it is solving the purpose for which that it is developed. This testing involves various types through which one can make sure that the software is reliable. The program was tested logically and pattern of execution of the program for a collection of data are repeated. Thus the code was exhaustively checked for all possible correct data and with an intent to search and find out whether the developed software met with the required requirements or not and to identify the defects and to

ensure that the product is defect free in order to produce a top quality product

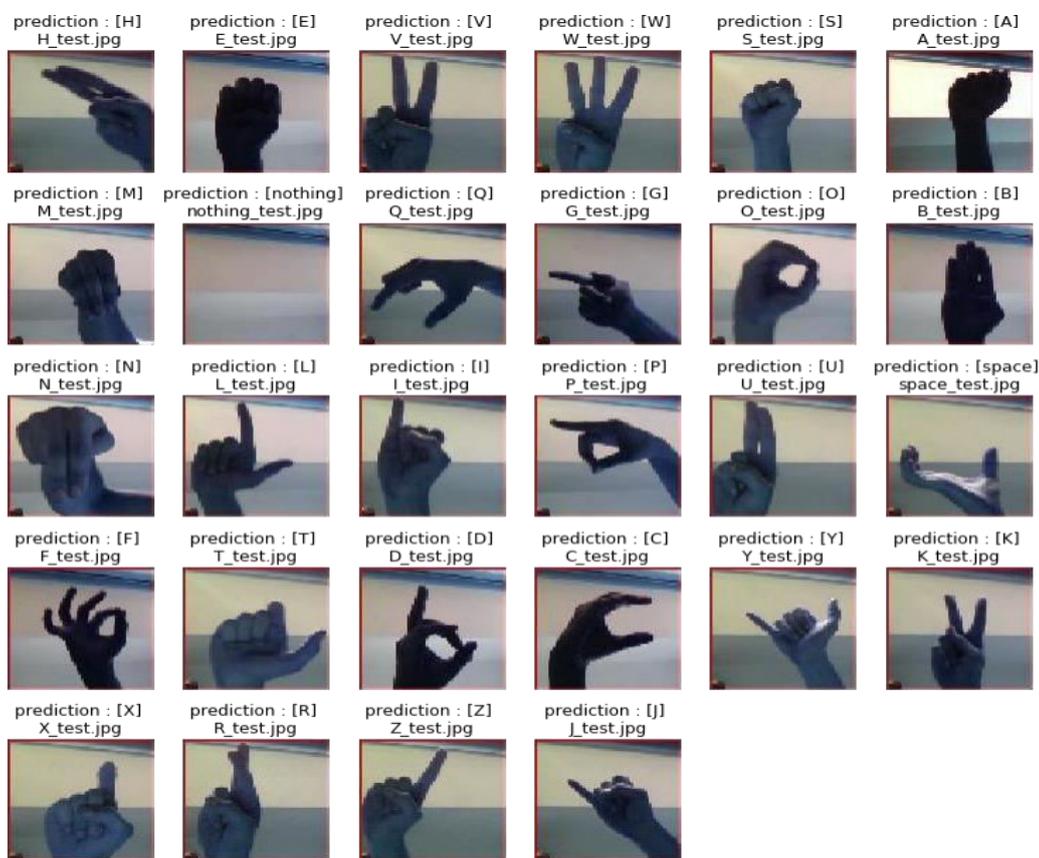
**Sign to Speech Outputs:**



**Speech to Sign Outputs :**



**Test Output:**



**REFERENCES**

1. Jangong Han, Enhance Computer Visions with Microsoft Kinects Sensor: A Review, IEEE TRANSACTION ON CYBERNETICS.
2. Microsoft Kinects SDK, <http://www.microsoft.com/ens-us/kinectsforswindows/>.
3. Gunasekaran. K, Manikandan. R, International Journal of Engineering and Technology (IJET): Sign Language to Speech Using PIC Microcontroller.
4. RiniAkmeliawati, Melanie PO-LeenOoi et al, "Real-Time Malaysian Sign Language Translation using Color Segmentation and Neural Network. Instrumentation and Measurement Technology Conference Warsaw", Poland.IEEE.1-6,2007.
5. Yang quan,"Chinese Sign Language Recognition Based On Video Sequence Appearance Modeling",IEEE. 1537-1542,2010.
6. Wen Gao and GaolinFanga,"A Chinese sign language recognition system Journal of Pattern Recognition".2389-2402,2004.
7. Nicholas Born.,"Senior Project Sign Language Glove", California Polytechnic State University,1- 49,2010.
8. Kirsten Ellis and Jan Carlo Barca."Exploring Sensor Gloves for Teaching Children Sign Language. Advances in Human-Computer Interaction".
- 9.1- 8.,2012. TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 59, NO. 10,2695-2704, 2012. Kinect camera, <http://www.xbox.com/en-US/kinect/default.htm>.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details