IJIRCCE

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

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 3, March 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

# Sign Language Recognition Using Deep Learning

**T.Shalini [1] , G.Shiny [2] , G.Lakshmi Haritha [3] , Jasmitha.M [4] , Ch.Pushpalatha [5]**

Assistant Professor, Dept. of ECE, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India

UG Student, Dept. of ECE, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India

UG Student, Dept. of ECE, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India

UG Student, Dept. of ECE, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India

UG Student, Dept. of ECE, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India

**ABSTRACT:** Over time, sign language has developed into an amazing advancement. Regrettably, there are certain negative side effects associated with this language. Not everyone is adept at deciphering sign language when speaking with a deaf or mute individual. Without an interpreter, communication is challenging. A product that is adaptable and durable is required to address this. The sign language must be translated in order for it to be understood by the general public. The idea of a camera-based sign language recognition system that would be used to translate sign language motions to text is thereby to increase the participation of the deaf and the mute in communication. Major deep learning techniques, including the CNN algorithm, are available to identify hand motion or gesture and then translate the detected information into voice.

**KEYWORDS**: Deep Learning, CNN Algorithm, Camera, Interpreter

## I. INTRODUCTION

The most popular means of communication for those who cannot speak is sign language. It's a language where letters and words are expressed using hand gestures. Just in China alone, sign language users have been observed. It reaches 80 million in total, and particularly those who are unable to grasp sign language will always find it difficult to communicate with one another. The recognition of signs using the vision method has gained popularity during the past few decades. This is a system that utilises a camera to detect the data that was gathered through finger gestures. It is the most popular visual-based technique. The development of vision-based sign recognition systems has required a great deal of work on a global scale. Direct and indirect techniques can be used in vision-based gesture recognition systems. In the past, a technique based on eyesight was employed to recognise hand movements. Yet, this procedure has a significant environmental impact on the recognised image. And another disadvantage is they have to reveal their hands to in front of the camera. Compared to other industries, this one has received less attention. The communication gap between the unique person and the average person is one of this special person's main challenges. It is usually difficult for Deaf and Dumb people to converse with regular people. They feel uneasy because of this enormous struggle and prejudice in society.Pre-processing is then used to resize the images to their original dimensions.

Creating a system that can translate hand motions into text is the goal of this project. The goal of this project is to insert the images into the dataset and transform the images to text using database matching. The detection process involves watching hands move. The method produces text-based output that aids in bridging the communication gap between the deaf-mute and the general public.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

## II.RELATED WORK

In[1], This research paper represents hand motions corresponding to ISL English alphabets are captured using a webcam. In the captured frames, the hand is segmented, and the position of the fingers is used to recognise the alphabet. The angle formed between fingers, the number of fingers that are fully opened, fully closed, or semi-closed, and the identification of each finger are used for recognition.
In[2] the authors proposed an approach for improving the SLR system. They use sensors embedded in a glove to detect gestures and translate them into speech using a Bluetooth module and an Android Smart Phone. The gloves will aid in the production of artificial speech, making it more similar to daily communication, which is hard for those with hearing disabilities to achieve.
In[3], the authors proposed multiple background subtraction algorithms that are tested, and then one of them is selected for the further process of action recognition. The next step is feature extraction which requires extracting important features from the image frame. The proposed novel action recognition algorithm makes use of a small image block (DFT).

## III. METHODOLOGY

**Existing Method:**
**Glove-based method**: The signer must wear a hardware glove while the hand movements are being recorded in this glove-based method, putting into practice a hand gesture detection system based on data gathered by a data glove. To obtain the hand's orientation along three axes, an accelerometer sensor was affixed to the glove's upper surface.
With this knowledge, the system would start making decisions using artificial intelligence technologies, such as neural networks, to identify the American Sign Language (ASL) signs for fingerspelling.
Although the glove-based method has an accuracy of over 90%, it appears a little awkward for application in everyday situations.

**Proposed Solution:**
The fundamental goal is to create an intuitive, straightforward solution that makes it easier for most individuals to communicate with deaf and mute persons. Here, we try to identify the hand motions or signals first before displaying the appropriate word. The first step entails utilizing a webcam and a posture estimation library to capture the gesture. The camera records the image, which is then processed in the tensor-flow tool to create text.
The steps below can be used to complete the project of sign language recognition:
1. Gathering Images
2. Labeling the gathered pictures
3. Using the dataset to train.
4. Data prediction

Cameras are the primary input method for sign language recognition (SLR). The SLR input data is an easily capturable gesture image by the camera.
Image processing is a technique for applying various procedures to an image in order to improve it or extract some relevant information from it. It is a kind of signal processing where the input is an image and the output can either be another image or features or characteristics related to that image.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

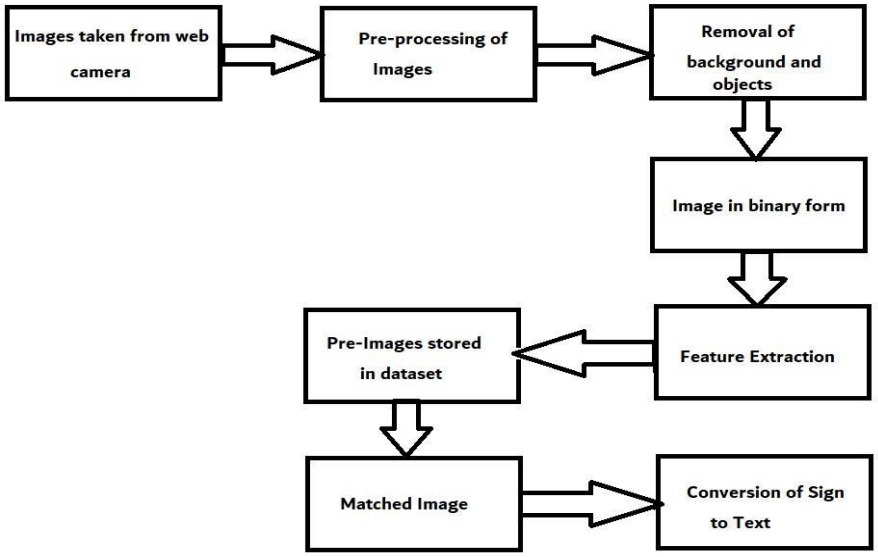


**Fig 1: Block diagram**

Initially, webcam images are captured, and pre-processing of the images is performed. The goal of this pre-processing is to enhance the quality of the images so that we can analyse them more effectively. Pre-processing allows us to reduce undesirable distortions and improve features that are important for the application we are working on. The removal of background and object is necessary because extraneous background objects frequently cause significant distortion during the image stitching process.As the image was read from the location, it is then transformed into binary form. Convert the colored image to its Grayscale form first because it is more straightforward and contains RGB layers.
Create a threshold mark such that any pixels above it change white and any pixels below it turn black. then do the feature Extraction Since that feature extraction is a step in the dimensionality reduction process, doing it this way is the most effective course of action. By choosing and combining variables into features, feature extraction helps to extract the best feature from those large data sets, effectively lowering the amount of data. These features are simple to use while still accurately and uniquely describing the real data set.The photos are now prepared for comparison, where they are compared to previously learned datasets that have been stored.If the image was successfully matched with the dataset then the output will be displayed in the form of text otherwise system stops recognizing the output.

## IV. IMPLEMENTATION

The bare minimum hardware needed to implement this is as follows:
1. Operating systems: Linux and Windows
2. Processor: Intel i3 at a minimum
3. Ram: 4 GB minimum
4. Hard disc: no less than 250 GB

The software requirements are listed here. Python Idle 3.7, Google Colab, or Anaconda can all be used to accomplish these tasks. Make sure Python or one of the other software mentioned above is installed on your system before starting the project. after installing python, import all the modules or libraries which are shown below.

**A.Tensorflow:**

A free and open-source software library called TensorFlow is used for differentiable programming and dataflow across a variety of activities. It is a symbolic math library that is also utilized by neural network applications in machine learning.

**B.Numpy**

A general-purpose array processing package is called Numpy. It offers a multidimensional array object with outstanding speed as well as capabilities for interacting with these arrays

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

**C.Pandas**

With its potent data structures, Pandas, an open-source Python library, offers high-performance data manipulation and analysis tools. Python was mostly utilized for data preprocessing and munging. Python and Pandas are utilized in a variety of academic and professional sectors, such as finance, economics, statistics, analytics, etc.

**D.Matplotlib**

Matplotlib is a Python 2D plotting library that produces publication-quality figures in a variety of hardcopy formats and interactive environments across platforms.

**E.Scikit – learn**

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

And there are some other basic modules that are to be imported they are time, uuid, etc…We are using deep learning Convolution Neural Network to train hand gesture photos and then this trained model can be used to predict those trained hand gestures from webcam.

**CNN Algorithm:**

Convolutional neural networks, often known as CNNs or ConvNets, are particularly skilled at processing input with a grid-like design, such as images. A digital image is a binary representation of visualdata.

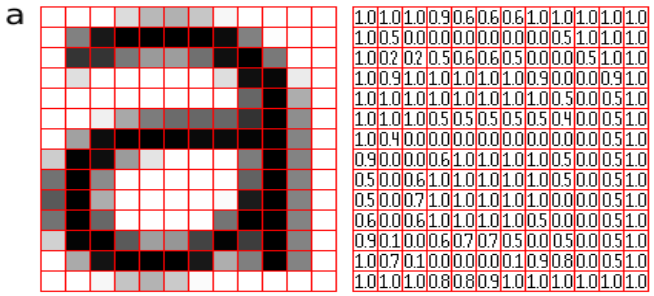


**Fig 2: Binary representation of data**

Once the images are being trained with this algorithm, The following modules are displayed on the screen in the process of presenting the output. They are as mentioned below

- Upload Hand Gesture Dataset
- Preprocess dataset
- Model Generation
- Train CNN Gesture Images
- Sign Language Recognition from Webcam
- Extract image from webcam
- Convert image to binary or grey format and back ground removal
- Extract features from image

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

## V.RESULTS

**1.**To run the project double click on the 'run.bat' file to get the below screen

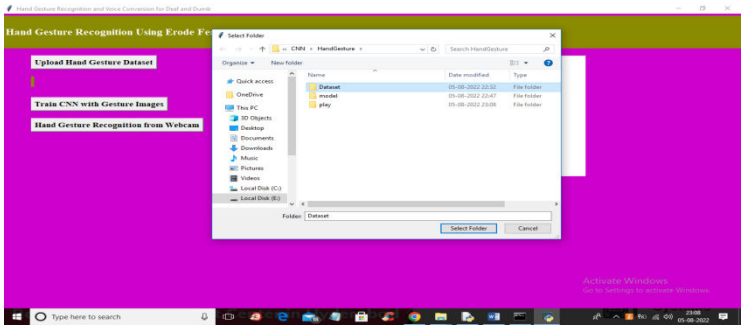


**Fig 4: Display Screen**

2. In above screen clicking on 'Upload Hand Gesture Dataset' button to upload dataset and then selecting and uploading Dataset folder to get below screen

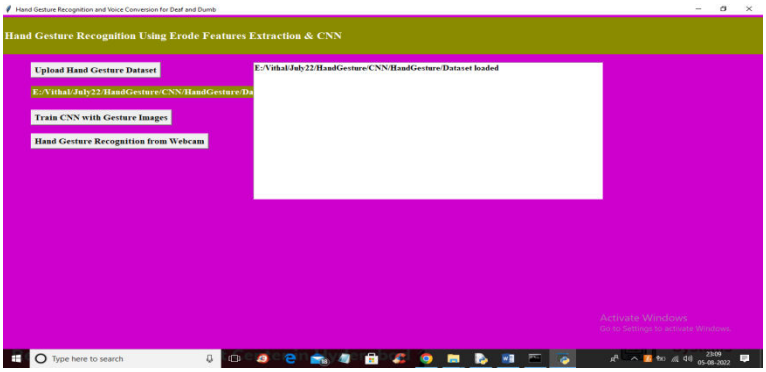


Fig 5: Upload hand gesture dataset

3. In above screen dataset loaded and now click on 'Train CNN with Gesture Images' button to train CNN and get below output

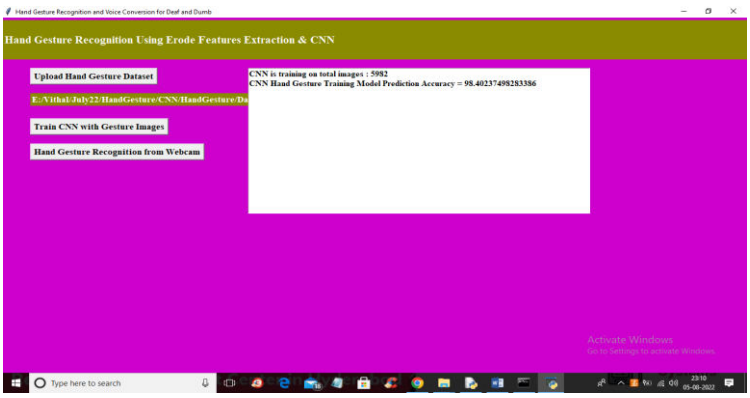


Fig 6: Train images with CNN algorithm

4. In above screen training is completed and we got CNN accuracy as 98% and now click on 'Hand Gesture Recognition from webcam'.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

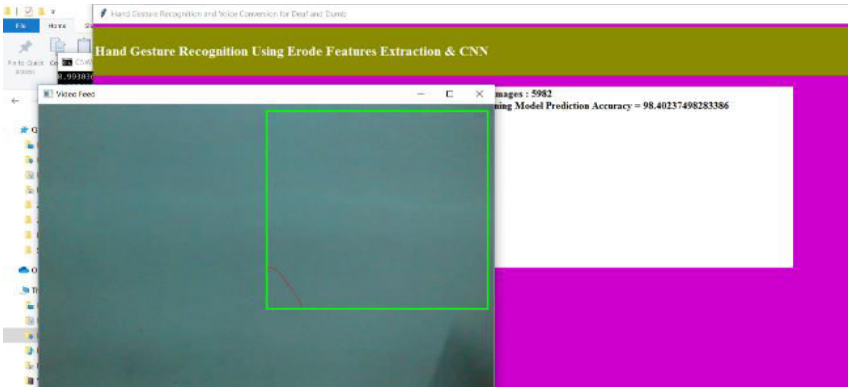


**Fig 7: Image collection frame**

5.Now show the symbol of HELLO in the green coloured rectangular box and the output will be displayed in the form of yellow coloured text showing hello.

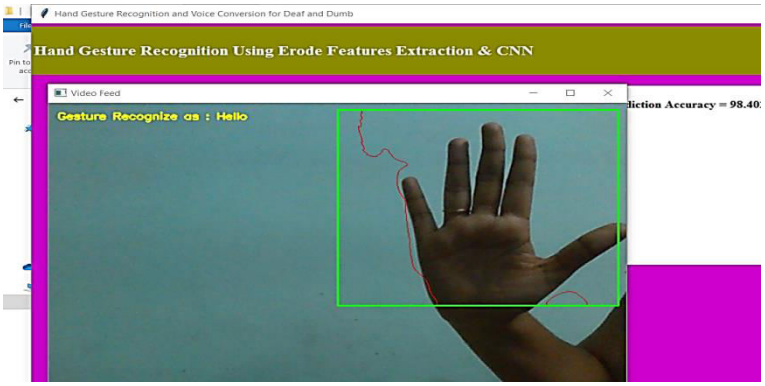


**Fig 8: Recognition of HELLO**

6.Now let us do the same for the sign STOP and the output will be displayed as shown

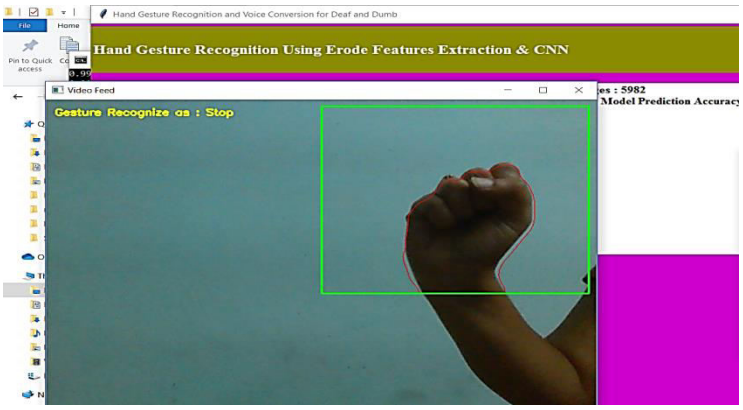


**Fig 9: Recognition of STOP**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

7. The below picture represents the output of the RIGHT

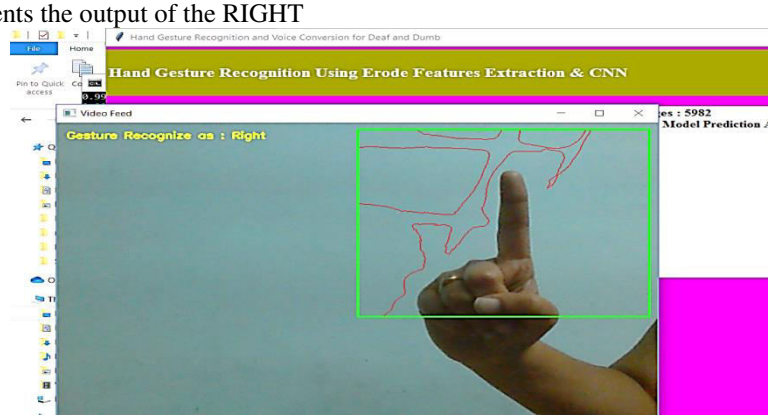


**Fig 10 : Recognition of RIGHT**

8.Let us do the same for BACKWARD and the output is displayed as shown below

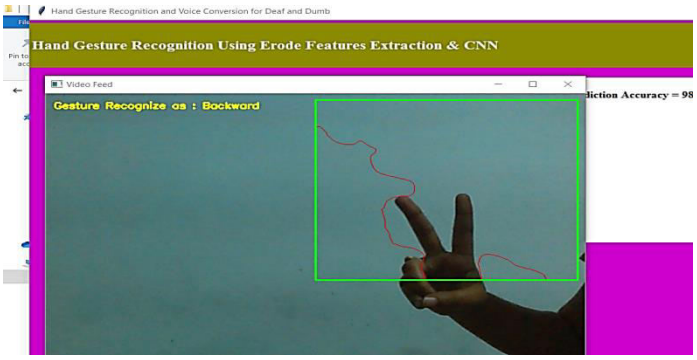


**Fig 11 : Recognition of Backward**

## VI. CONCLUSION& FUTURE SCOPE

Convolutional neural networks, a deep learning technique, and image processing were used to recognise hands for deaf and mute people. The technique uses a webcam to capture photos of hand gestures as input, and it outputs text. This system's implementation provides accuracy of up to 90% and functions well in the majority of test cases. It would even be possible to expand our dataset in the future and interpret more gestures into voice. Consequently, even those who are deaf or mute will be able to converse with ease utilising this technique.This technique for recognising hand gestures effectively solves the issue of processing and extracting video frames. Different hand gestures can be identified and used as computer input in the future. The hand motions used to represent numbers can also be translated into orders that will carry out relevant actions immediately. Further work can be done to improve the recognition capacity for diverse lighting circumstances, which is a problem in this project**.**

## REFERENCES

[1] Shangeetha, R. K., V. Valliammai, and S. Padmavathi. "Computer vision-based approach for Indian Sign Language character recognition." Machine Vision and Image Processing (MVIP), 2012 International Conference on. IEEE, 2012.

[2] Heera, S.Y., Murthy, M.K., Sravanti, V.S. and Salvi, S., 2017, February. Talking hands—An Indian sign language to speech translating gloves. In 2017 International conference on innovative mechanisms for industry applications (ICIMIA) (pp. 746-751). IEEE.

[3] Kumari, Sonal, and Suman K. Mitra. "Human action recognition using DFT." Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG), 2011 Third National Conference on. IEEE, 2011.[4] M. Wald, "Captioning for Deaf and Hard of Hearing People by Editing Automatic Speech Recognition in Real Time", Proceedings of 10th International Conference on Computers Helping People with Special Needs ICCHP 2006, LNCS 4061, pp. 683- 690.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103031 |

[5] A. Y. Satpute, A. D. Bhoi, and T. Engineering, "ELECTRONIC SPEAKING SYSTEM FOR DUMB," vo!. 6, no. 3, pp. 1132-1139, 2013.
[6] R. R. Itkarkar and A. V. Nandi, "Hand gesture to speech conversion using Matlab," in 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013, pp. 1-4.
[7] Jingdong Zhao, Li Jiang, Shicai Shi, Hegao Cai, Hong Liu, G. Hirzinger, "A Five-fingered Underactuated Prosthetic Hand System",Proceedings of the 2006
[8] J. Gangrade and J. Bharti, "Vision-based Hand Gesture Recognition for Indian Sign Language Using Convolution Neural Network," IETE Journal of Research, vol. 31, pp. 1-10, 2020.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details