IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Sign Language Translator Using Deep Learning

N. Greeshma[1], Md. Abdul Samad[2], K. Mukesh[3], Bibi Jainab[4]

[1]Graduation Student, Bachelor of Technology, Department of Information Technology, Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada, India

[2]Graduation Student, Bachelor of Technology, Department of Information Technology, Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada, India

[3]Graduation Student, Bachelor of Technology, Department of Information Technology, Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada, India

[4]Graduation Student, Bachelor of Technology, Department of Information Technology, Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada, India

Abstract - Everyone can communicate effectively with hand gestures. Hand gesture recognition, as opposed to the traditional text-based or GUI-based method, can be viewed as a means for computers to better comprehend human body language, resulting in a fuller interaction between humans and computers. Hand gestures can simplify work by sending orders to the computer using simple hand movements. The camera captures the movements, the computer interprets them, and suitable action is conducted in response to the gesture. This project is built on sign language translation, which captures hand movements and converts them into commands that the computer can comprehend and give the output as audio and text format and do relevant tasks. In this project, we first use a camera to capture the gesture. The collected picture is then pre-processed for feature extraction, and a convolutional neural network is implemented to successfully detect motions and train the machine.

***Keywords*: Hand Gestures, Image Processing, Deep Learning**

I. INTRODUCTION

More than 70 million individuals are deaf or hard of hearing. Sign languages allow these people to communicate using postures, body movements, eyes, brows, and hand gestures. Sign language is a natural means of communication for those who have hearing and/or speech problems. Signs are motions done with one or both hands, accompanied by certain face expressions. Although deaf, hard-of-hearing, and mute persons may readily interact with one another, integration in educational, social, and professional situations is a substantial challenge for the differently abled.

There is a communication barrier between an unimpaired person who is unaware of the sign language system and an impaired person who desires to communicate. However, there is no one standard form of sign language, and it differs from area to region. People communicate with one another through sharing their ideas, opinions, and experiences. Hence, in addition to utilizing traditional sign languages, every deaf or dumb person can have signs that reflect his or her lifestyle, habits, and traditions. Deaf individuals have difficulty communicating with hearing people through writing. Most hearing individuals struggle to grasp special gestures used by Deaf persons when writing. Paper and pen are insufficient for efficient communication.

In recent years, computers have made tremendous advances in recognizing and rendering sign language. Because of technology breakthroughs, the way the world operates is continuously changing and improving. As programs have evolved over the previous two decades, barriers for the differently abled have dissipated. Image processing, artificial intelligence, and pattern matching technologies are being developed by researchers to assist these individuals engage and learn. The project's purpose is to assist individuals in overcoming these challenges by creating vision-based technology for identifying and converting sign language into text. This project combines motion detection and gesture recognition in real time. A precise gesture must be performed by the user. The camera catches and identifies the gesture, and the correct depiction is displayed.

II. RELATED WORK

Paper [1], the authors used flex sensors and IMU to recognize their sign symbol and voice output. These are interfaced with microcontroller, which programmed to obtain corresponding output.
In this the drawbacks are:
Measurement errors caused by moving garments.
Electromagnetic Interference.
External pressure sensitivity.

Paper [2], the authors proposed a kind of Attention-Based network utilizing the sampling and skeletal feature to improve the accuracy of SLR. Later they adopted a temporal Attention-Based BLSTM for distributing weights to the features of different keyframes.
Drawback: Feature extraction and Matching algorithm need to develop the real-time recognition system.

Paper [3], the author describes the development of a system that enables editors to correct errors in the captions as they are created by Automatic Speech Recognition. This offers the potential to provide automatic real-time captioning to follow the captions and transcript than to follow the speech of the lecturers.

Paper [4], the author uses a camera for capturing the gesture which is followed by all the Image processing techniques and algorithms and gives an output. But the accuracy may not be efficient when compared to present algorithms.

III. PROBLEM STATEMENT

Dumb individuals communicate through hand signals, therefore normal people have difficulty comprehending their language by signs created. As a result, technologies that identify diverse signals and deliver information to ordinary people are required.

IV. AIM

Our project's goal is to create a concept of a virtual chatting system without sensors for individuals in need, with this idea reaching success through image processing and human interaction input via hand gestures. This mostly benefits those who are unable to communicate with others.

V. TECHNOLOGIES USED

OpenCV

OpenCV stands for Open-Source Computer Vision Library. It has C++, Python and Java interfaces and supports many operating systems like Windows, Linux, Mac OS, iOS and Android. OpenCV is designed to improve computational efficiency and provide a strong focus on real-time applications.

MediaPipe

MediaPipe is Google's open-source framework, used for media processing. It is cross-platform or we can say it is platform friendly. It can run on Android, iOS, and the web that's what Cross-platform means, to run everywhere. MediaPipe is a customizable machine learning solutions framework developed by Google. It is an open-source and cross-platform framework, and it is very lightweight. MediaPipe comes with some pre-trained solutions such as face detection, pose estimation, hand recognition, object detection etc.

Python

Python is an interpreted, high-level, general-purpose programming language.

PIL (Python Imaging Library)

Python Imaging Library which is also known as Pillow in newer versions is a free library for Python. It provides support for displaying and manipulating, different image file formats.

Numpy

NumPy is a library used by the Python programming language. It provides support for large, multidimensional arrays and matrices, and also has a large collection of high- level mathematical functions to operate on these arrays.

Tensorflow

TensorFlow is an open-source software library which provides high performance numerical computation. It provides strong support for machine learning and deep learning and the flexible numerical computation core is used in many other scientific domains.

Keras

It is an open-source neural-network library written in Python language. It is capable of running on top of TensorFlow and it has many implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and many tools to simplify working with image and text data also available when necessary.

VI. PROPOSED METHOD

In the proposed system, the incapable or dumb person should give the system with a gesture or sign image. The system uses the image processing approach to analyse the sign input and classifies it to the known identity. When the input image matches the specified dataset, it then starts the speech media through the system. In addition, the output will be shown in text format.

This is a working prototype for translating sign language to speech and text. The goal of this study is to offer a society-wide application that will facilitate communication between deaf and dumb persons.

VII. METHODOLOGY

In this section, we elaborate on our hierarchical architecture that enables the-state-of-the-art CNN models to be used in real-time gesture recognition applications as efficiently as possible. After introducing the architecture, training details are described. Finally, we give a detailed explanation for the used post processing strategies that allow us to have single-time activation per gesture in real-time.

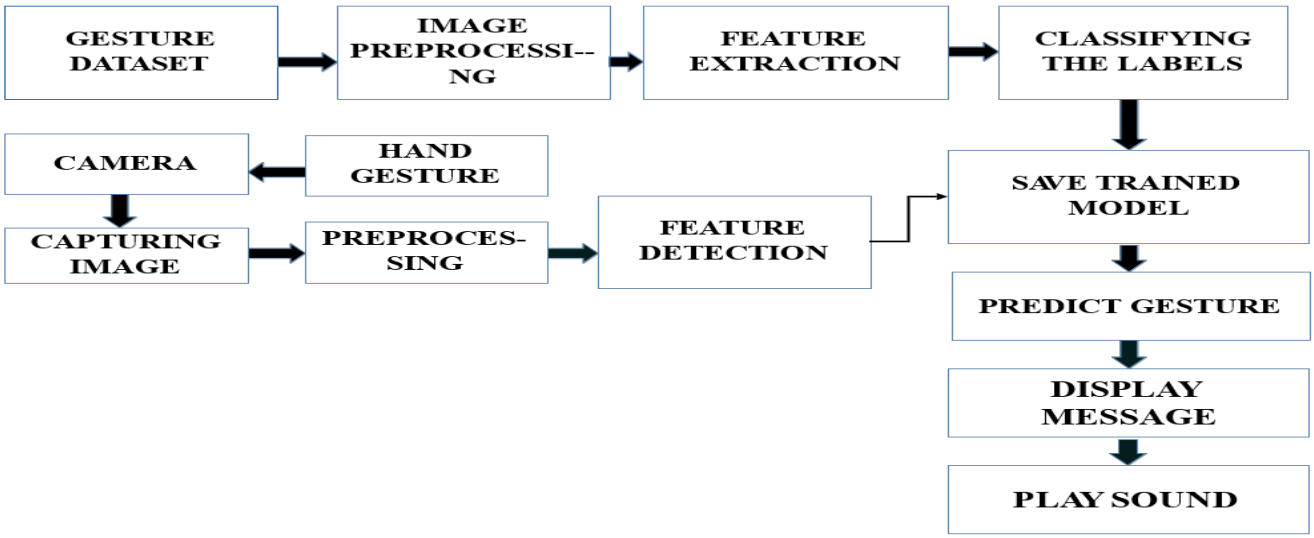

Figure 1: Block Diagram

CONVOLUTION NEURAL NETWORK

Our project uses Convolution Neural Network to anticipate hand motions. This network operates on a Tensorflow backdrop, with Keras as the primary supporting library for neural network training and prediction. A distinguishing aspect of our Prediction model is that when a new Gesture is required, we may add it without having to import a new Gesture Dataset for Training the Network, because we can build Dataset and use it for Training the neural network.

After performing Image Processing on the collected photos, such as edge detection, hand contour analysis, and so on. The neural network is then fed the images. To decrease the temporal complexity of the Training and Prediction phases, each picture that is sent to the Neural Network is 100*100 pixels in size. A 100*100-pixel RGB picture is presented as input to the Neural Network. This approach employs four levels, including one input layer, two hidden layers, and one output layer.

VIII. RESULTS AND DISCUSSIONS

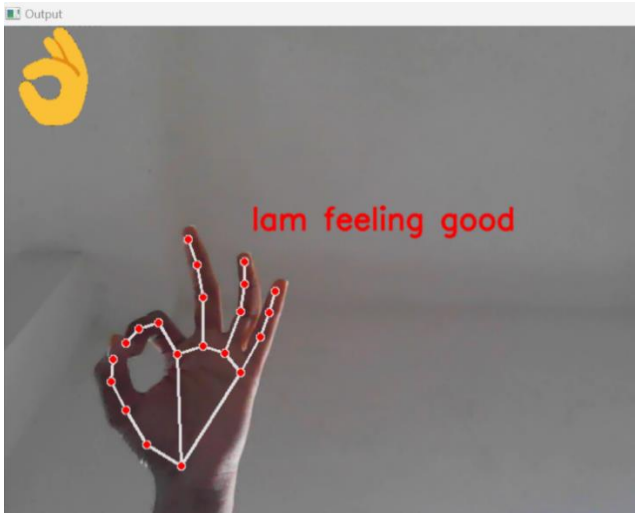

Figure 2: Output “I am feeling good”

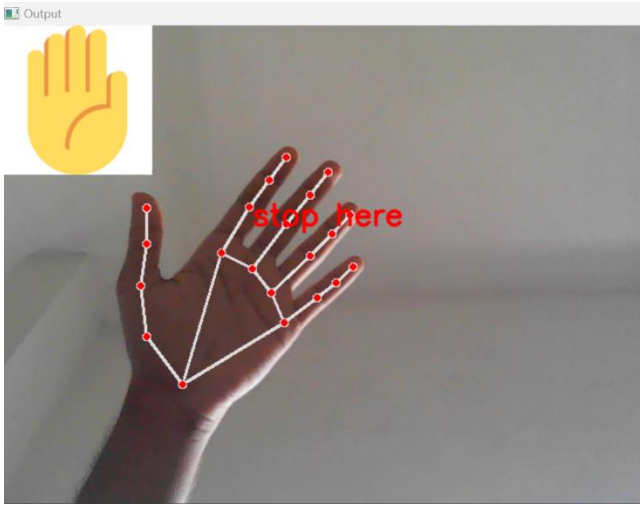

Figure 3: Output “Stop here”

Figure 4: Output “Give me some food”

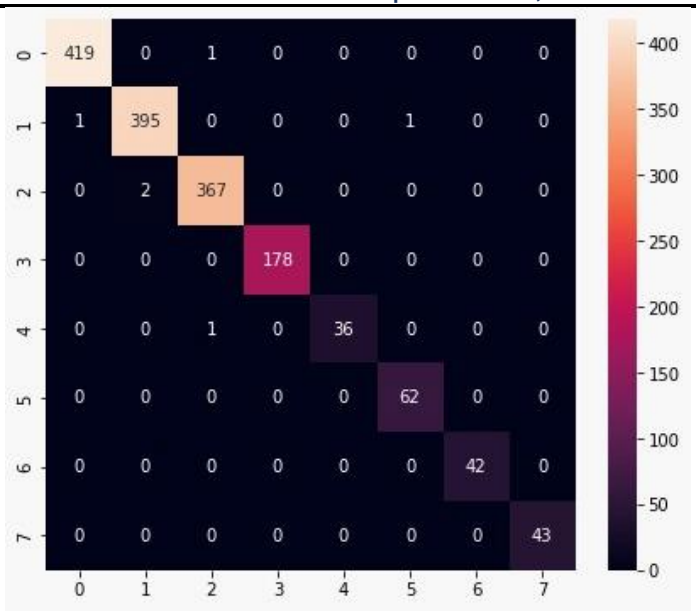

Figure 5: Confusion Matrix

IX. CONCLUSION

In our system, we used several image processing techniques to fully isolate the hand from the backdrop and feed it into the deep neural network to anticipate motions. By imposing appropriate limits, these defined approaches may be employed precisely and efficiently.

X. ACKNOWLEDGMENT

We would like to thank our guide Assistant Professor J.V.N Raju and our Information Technology Department for their support and guidance during the implementation of this project and also while writing this paper.

REFERENCES

1. https://www.researchgate.net/publication/355858241_Sign_Language_Translator_Using_Machine_Learning

2. G. Anantha Rao; K. Syamala; P. V. V. Kishore; A. S. C. S. Sastry, “Deep Convolutional Neural Networks for Sign Language Recognition”, IEEE, 2018.

3. Kshitij Bantupalli, Ying Xie, “American Sign Language Recognition using Deep Learning and Computer Vision”, IEEE,2018.

4. Sang-Ki Ko, Jae Gi Son, Hyedong Jung, “Sign Language Recognition with Recurrent Neural Network using Human Keypoint Detection”, Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems, October 2018.

5. Md. Moklesur Rahman, Md. Shafiqul Islam, Md. Hafizur Rahman, Roberto Sassi, “A New Benchmark on American Sign Language Recognition using Convolutional Neural Network”, IEEE,2019.

6. Learning and Computing, Vol. 9, No. 6, December 2019.

7. Neena Aloysius, M Geetha, “Understanding Vision Based Continuous Sign Language Recognition”, 17 May, 2020.

8. K Amrutha, P Prabu, “ML Based Sign Language Recognition System”, IEEE, 2021.

9. Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B, “Static Sign Language Recognition Using Deep Learning”, International Journal of Machine.