

# **REAL TIME HAND SIGN DETECTION (USING KERAS CNN)**

SHREERAM.S<sup>1</sup>, Utkarsh Sharma<sup>2</sup>, Bassar Patel<sup>3</sup>

Rajalakshmi Engineering College, Thandalam, Chennai-602105 <sup>1</sup>, IBM India Private

Limited, Bengaluru, Karnataka 560071<sup>2 3</sup>

{200701234@rajalakshmi.edu.in<sup>1</sup>,sshreeram0511@gmail.com<sup>1</sup>

utkarsh@edunetfoundation.org <sup>2</sup>, bassar@edunetfoundation.org <sup>3</sup>}

## ABSTRACT

Communication is defined as the act of sharing or exchanging information, ideas or feelings. To establish communication between two people, both are required to have knowledge and understanding of a common language. But in the case of deaf and mute people, the means of communication are different. To overcome their barrier, one can build a model based on machine learning. A model can be trained to recognize different gestures of sign language and translate them into English. This will help a lot of people in communicating and conversing with deaf and mute people. A method proposed train a TensorFlow model to create a real-time Sign Language Recognition system. The system achieves a good level of accuracy even with a limited size dataset.

**Keywords:** Sign Language Recognition (SLR), Computer Vision, Machine Learning, Indian Sign Language.

## INTRODUCTION

Communication can be defined as the act of transferring information from one place, person, or group to another. It consists of three components: the speaker, the message that is being communicated, and the listener. It can be considered successful only when whatever message the speaker is trying to convey is received and understood by the listener. It can be divided into different categories as follows : formal and informal communication, oral (face-to-face and distance) and written communication, non-verbal, grapevine, feedback, and visual communication, and the active listening.

The formal communication (official communication) is steered through the channels that are pre-determined. The unofficial or grapevine communication is the spontaneous communication between individuals in one's profession that does not have any formal protocol or structure.

The oral communication (face-to-face and distance) is the communication in which words are exchanged between people who are present in front or at a distance (with the help of technology including voice and video calls, webinars, etc.). The written communication is the communication in which letters, emails, notices, or any other written form is used for communicating. The non-verbal communication is the communication that uses gestures, facial expressions, body language, etc.

The feedback communication happens when a person gives feedback on some product or service provided by an individual or a company. The visual communication occurs when a person gets information from a visual source like televisions, social networking, or any other source. Active listening is when a person listens to and understands what the other individual is trying to convey so that the communication becomes more meaningful and effective.

Non-verbal communication helps deaf and dumb people to communicate amongst themselves and with others. Deaf is a disability that impairs a person's hearing ability and makes them incapable to hear while dumb is a disability that impairs the speaking ability and makes them incapable to speak. Not being able to speak or listen makes it difficult to establish communication with others. This is where sign languages come into the role, it enables a person to communicate without words. But a problem still exists, not many people possess the knowledge of sign language. Deaf and dumb may be able to communicate amongst themselves using sign languages but it is still difficult for them to communicate with people having normal hearing and vice-versa due to the lack of knowledge of sign languages. This issue can be resolved using a technology-driven solution. By using such a solution, one can easily translate the gestures of sign language into the commonly spoken language, English.

A lot of research has been done in this field and there is still a need for further research. For gesture translation, data gloves, motion capturing systems, or sensors have been used. Vision-based SLR systems have also been developed previously. The existing Indian Sign Language Recognition system was developed using machine learning algorithms with MATLAB. Authors have worked on single-handed and double-handed gestures.

They used two algorithms to train their system, K Nearest Neighbors Algorithm and Back Propagation Algorithm. Their system achieved 93-96% accuracy. Though being highly accurate, it is not a real-time SLR system. The objective of this paper is to develop a real-time SLR system using TensorFlow object detection API and train it using a dataset that will be created using a webcam.

The rest of this paper after the introduction is organized as follows. Section 2 presents the related work on the SLR system. Section 3 describes the data acquisition and generation. Section 4 focuses on the methodology of the developed system. Section 5 presents the experimental evaluation of the system, and finally, Section 6 concludes the paper with future work.

## BLOCK DIAGRAM

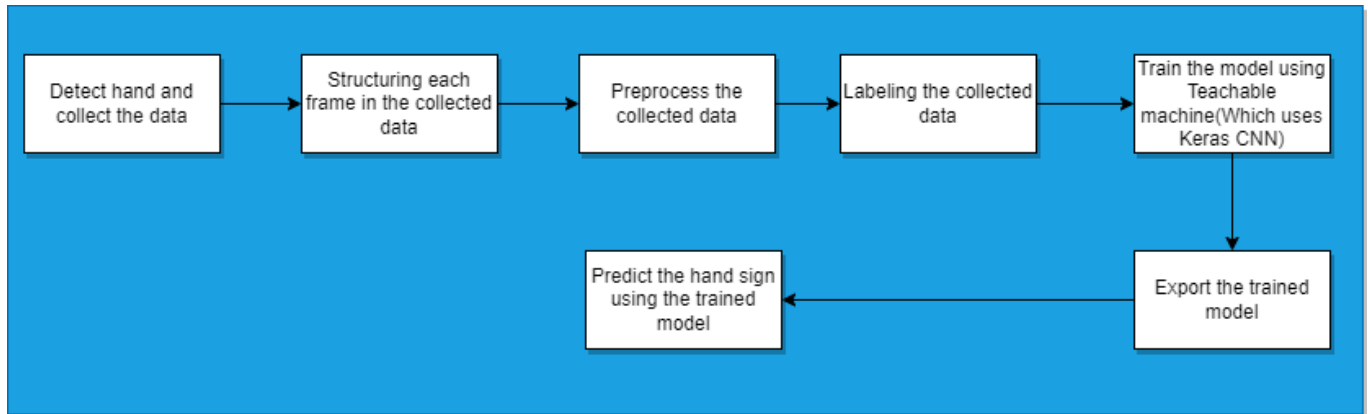


Fig 1.1 Block Diagram

## TECHNOLOGIES

### CONVOLUTIONAL NEURAL NETWORK (CNN)

Deep learning is a very significant subset of machine learning because of its high performance across various domains. Convolutional Neural Network (CNN) is a powerful image processing deep learning type often using in computer vision that comprises an image and video recognition along with a recommender system and natural language processing (NLP).

CNN uses a multilayer system consists of the input layer, output layer, and a hidden layer that comprises multiple convolutional layers, pooling layers, fully connected layers.

### PLATFORM AND PROGRAMMING LANGUAGE

Jupyter notebooks basically provides an interactive computational environment for developing Python based Data Science applications. They are formerly known as ipython notebooks. The following are some of the features of Jupyter notebooks that makes it one of the best components of Python ML ecosystem –

- Jupiter notebooks can illustrate the analysis process step by step by arranging the stuff like code, images, text, output etc. in a step by step manner.
- It helps a data scientist to document the thought process while developing the analysis process.
- One can also capture the result as the part of the notebook.
- With the help of Jupyter notebooks, we can share our work with a peer also.

Python is a programming language that is preferred for programming due to its vast features, applicability, and simplicity. The Python programming language best fits machine learning due to its independent platform and its popularity in the programming community.

### **DATA SET**

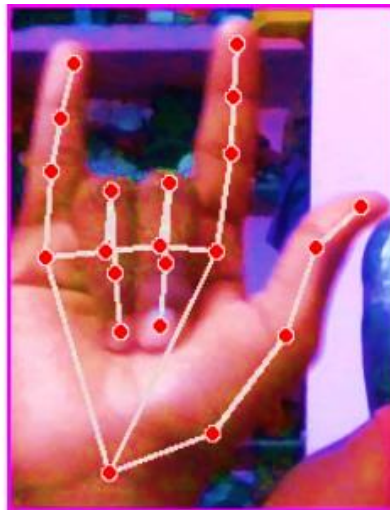


Fig 1.2- I Love You Sign - Dataset



Fig 1.3- No - Dataset

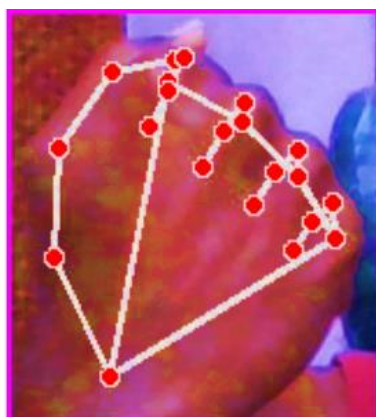


Fig 1.4- Sorry - Dataset

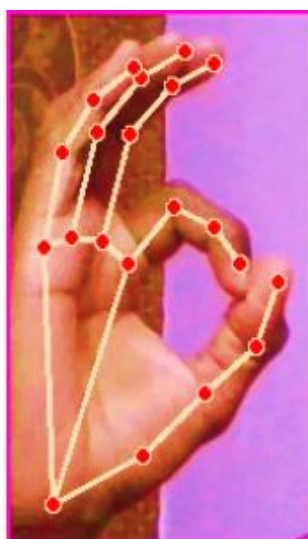


Fig 1.5- Super – Dataset

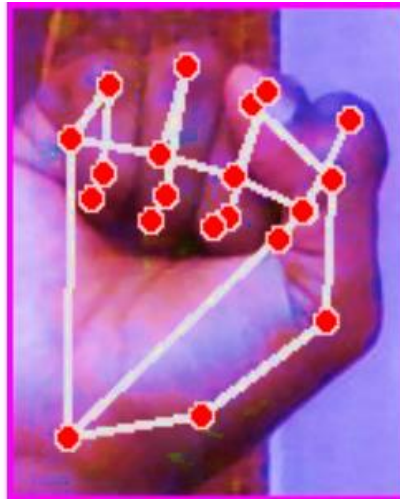


Fig 1.6- Yes – Dataset

## RESULTS

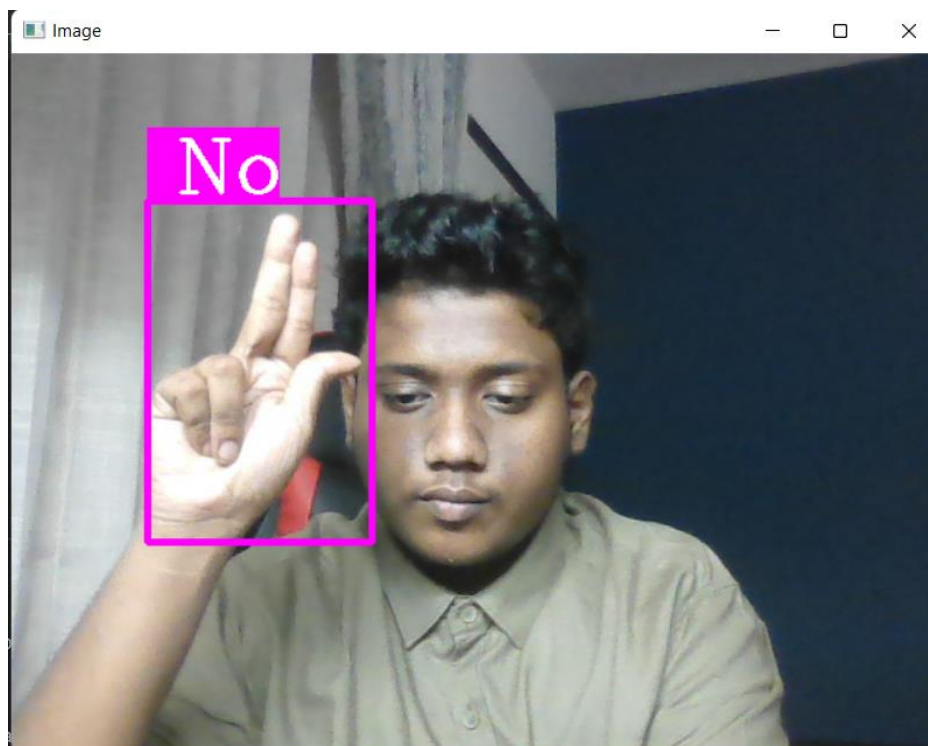


Fig 1.7- No – Result

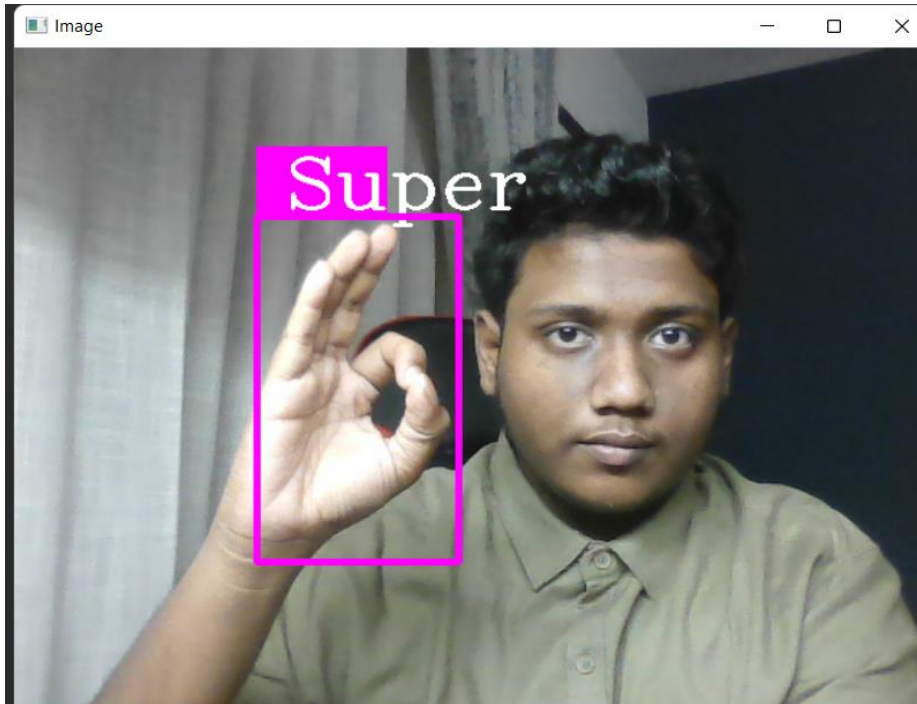


Fig 1.8- Super – Result

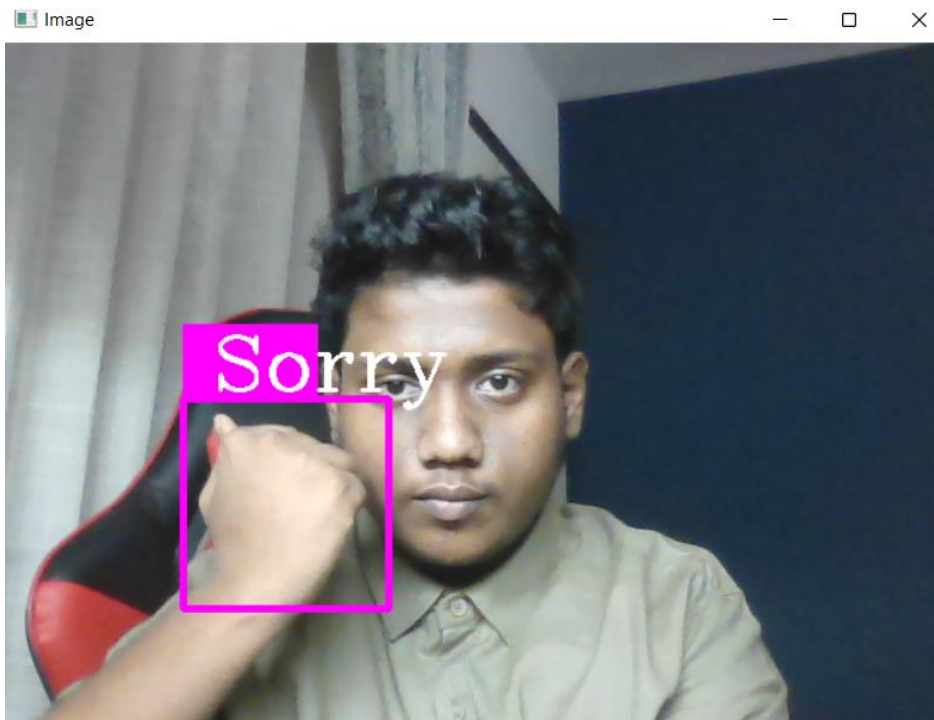


Fig 1.9- Sorry – Result





Fig 1.10- Yes – Result

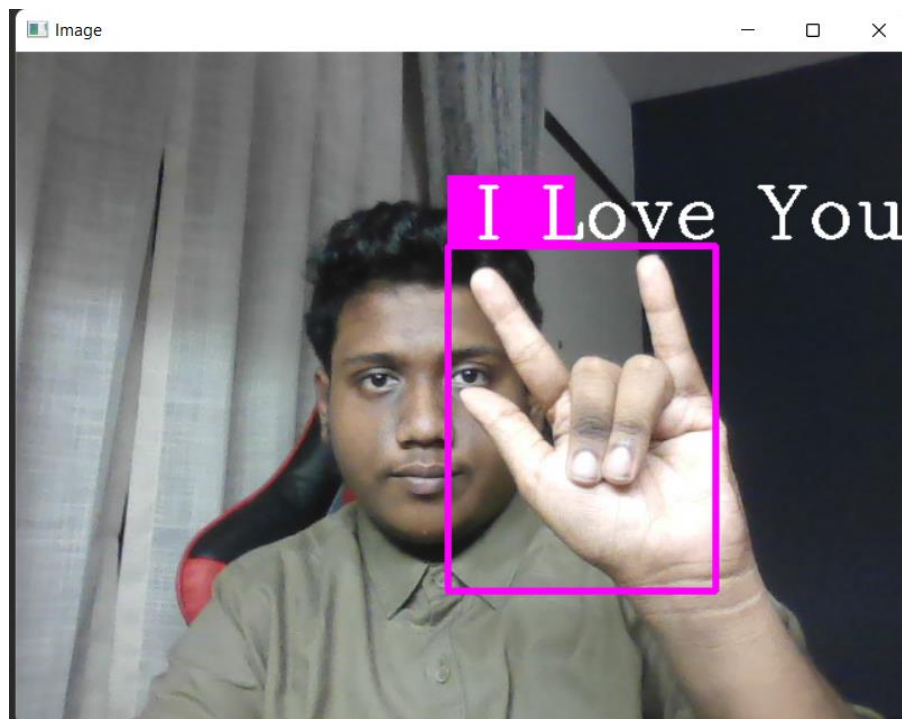


Fig 1.11- I Love You – Result

CNN with Keras (Accuracy = 98%)

## **CONCLUSION**

Nowadays, applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted to perform various applications. When an image is transformed from one form to another such as digitizing, scanning, and communicating, storing, etc. degradation occurs. Therefore, the output image must undertake a process called image enhancement, which contains of a group of methods that seek to develop the visual presence of an image. Image enhancement is fundamentally enlightening the interpretability or awareness of information in images for human listeners and providing better input for other automatic image processing systems. Image then undergoes feature extraction using various methods to make the image more readable by the computer. Sign language recognition system is a powerful tool to prepare an expert knowledge, edge detect and the combination of inaccurate information from different sources. The intend of convolution neural network is to get the appropriate classification. The IBM Internship 2022 had enhanced my technical skills on the domain of artificial intelligence and via this opportunity I was able to explore lot about the machine learning to acquire the domain knowledge.

## **FUTURE SCOPE**

The proposed sign language recognition system used to recognize sign language letters can be further extended to recognize gestures facial expressions. Instead of displaying letter labels it will be more appropriate to display sentences as more appropriate translation of language. This also increases readability. scope of different sign languages can be increased. More training data can be added to detect the letter with more accuracy. This project can further be extended to convert the signs to speech.

## REFERENCES

1. Kapur, R.: The Types of Communication. MIJ. 6, (2020).
2. Suharjito, Anderson, R., Wiryana, F., Ariesta, M.C., Kusuma, G.P.: Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input- Process-Output. *Procedia Comput. Sci.* 116, 441–448 (2017). <https://doi.org/10.1016/J.PROCS.2017.10.028>.
3. Konstantinidis, D., Dimitropoulos, K., Daras, P.: Sign language recognition based on hand and body skeletal data. *3DTV-Conference*. 2018-June, (2018). <https://doi.org/10.1109/3DTV.2018.8478467>.
4. Dutta, K.K., Bellary, S.A.S.: Machine Learning Techniques for Indian Sign Language Recognition. *Int. Conf. Curr. Trends Comput. Electr. Electron. Commun. CTCEEC* 2017. 333–336 (2018). <https://doi.org/10.1109/CTCEEC.2017.8454988>.
5. Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Verhoef, T., Vogler, C., Morris, M.R.: Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. *21st Int. ACM SIGACCESS Conf. Comput. Access.* (2019). <https://doi.org/10.1145/3308561>.
6. Rosero-Montalvo, P.D., Godoy-Trujillo, P., Flores-Bosmediano, E., Carrascal-Garcia, J.
7. Otero-Potosi, S., Benitez-Pereira, H., Peluffo-Ordonez, D.H.: Sign Language Recognition Based on Intelligent Glove Using Machine Learning Techniques. *2018 IEEE 3rd Ecuador Tech. Chapters Meet. ETCM 2018.* (2018). <https://doi.org/10.1109/ETCM.2018.8580268>.
8. Nikam, A.S., Ambekar, A.G.: Sign language recognition using image based hand gesture recognition techniques. *Proc. 2016 Online Int. Conf. Green Eng. Technol. IC-GET 2016.* (2017). <https://doi.org/10.1109/GET.2016.7916786>.