



# Sign Language Recognition System using Transfer Learning Technique

Shivani Y, Jarul R, Vandana N

Department of Electronics and Communication Engineering  
shivaniydv9@gmail.com, jarul.rana@gmail.com, vandananiranjana@igdtuw.ac.in  
Indira Gandhi Delhi Technical University for Women, Delhi, India

## Abstract

In this research paper, we have proposed a design which can be used to detect different static hand gestures using transfer Learning, python and OpenCv in real-time. Common words used in daily conversations such as - Hello, Yes, No, Thank you and Bye are translated into sign language and are successfully recognized by our system. The major steps involved in designing the system are - creation of our own dataset by capturing images with openCv and webcam, labelling each captured gesture for object detection, training and testing the dataset with transfer learning using SSD MobileNet, and finally, determining the hand gestures in real-time. The complete project is implemented on Jupyter notebook using python programming language.

**Keywords:** Transfer Learning, OpenCv, Tensorflow Object Detection, Hand gesture recognition

## 1. Introduction

With the advancement in technology, emerging fields such as Artificial Intelligence and Deep learning have been explored upto a greater extent. Communication is an essential part for performing daily activities. According to the World Health Organization (WHO), over 5% of the total world's population is suffering from hearing impairments and muteness. A significant portion of the world is deprived of communicating with other people effectively. Many sign languages are present today, they have been a vital support to differently abled community to interact with each other. But every normal person doesn't understand these sign languages. So, they require to first learn the language in order to express his/her thoughts with impaired ones. Hence, the gap of communication with each other still persist in the society. A lot of efforts have been made to bridge the gap in between differently abled people and normal ones because

everyone has equal rights to access the care, support, education, medical facilities and social interactions.

Therefore, sign language system helps to eradicate all the barriers that inhibit the deaf, dumb people to communicate and express their thoughts and ideas to the world. We have used common words such as – Yes, No, Thank you, Bye and Hello and translated them into signs which best represents the respective word. The idea to incorporate the words directly for detection instead of embedding the alphabets and numbers in our model is to make the system time efficient for communication. It will help to the impaired people to communicate with others at a faster rate.

Through this system, a person can express non verbally through hands movement swiftly rather wasting time in the forming each alphabet for each word. This minimizes the scope of miscommunication and misinterpretation. In similar manner, thousands of

words that we use on our daily basis can be assigned hand gestures and easily recognized.

In the past years, Sign Language recognition framework have been major focus of study. Many approaches are taken in regard, such as sensor based devices are developed, gloves based detection have been deployed [1]. But they require higher costs and complex computations. Our project focuses on building a static sign language detection system and its a vision based SLR architecture. That means, the system is able to determine the word signs made by the signer using openCv in real time. Therefore, it provides more accuracy and comes out to be more affordable model. The weak point of the system is the heavy usage of computing power while training down the model using Tensorflow Object Detection API.

## 2. Design and Implementation

Deep learning has shown tremendous growth in speech recognition and speech to text recognition [4]. Huge amounts of data is required to train deep learning models. Transfer learning is a technique to utilize a related task with a large amount of data available to help solve a specific task lacking sufficient data. Transfer learning is implemented by pre-training a network on the collected dataset. The pre training model – SSD MobileNet is used for transfer learning. Sign language recognition architecture has been divided into four segments, and they are as follows;

1. Creating our own dataset
2. Labelling captured images
3. Training the models using tensorflow object detection
4. Detecting the hand gesture

The proposed system can recognize the static word sign and represent its label for better communication. Initially, different hand gestures images were captured using our Web cam. Five words been translated into signs were collected and stored for training them later on, and they were – HELLO, ,THANKS, YES, NO, BYE.



Fig. 1 shows the hand gesture for YES



Fig. 2 shows the hand gesture for HELLO

Sign language recognition project is carried out using python programming language, OpenCv and Tensorflow

Object Detection software. The entire python code has been written in an open source, web based IDE called Jupyter notebook. The images are captured in bright and dim lighting, at different angles which would make greatly affect the accuracy of the system. Different intensity of light in the room, variation in pose, location adds to the variability. Fig.1-5 presents the words signs that will be fed onto the system.

The datasets are labelled using labelImg. A detection box is created around targeted hand gesture and provided with label. LabelImg is an open source, graphical image annotation tool especially used for labelling images for object detection. This will help us to detect the hand gestures by the specified label. The next step in the process is training and testing our model. For that, due to limited gathered dataset, we have divided 15 images of each label for training phase and remaining 5 images per label are put into testing phase. In order to train our models, we require to retain a huge amount of datasets for deep learning purpose. Hence, for our design, we have made transfer learning basis to do so.



Fig. 3 shows the hand gesture for BYE



Fig.4 shows the hand gesture for THANK YOU



Fig. 5 shows the hand gesture for NO

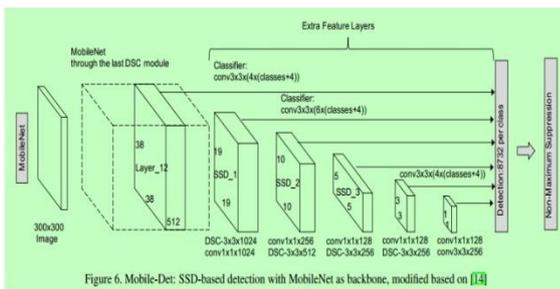


Fig. 6 shows the SSD MobileNet layered architecture

Tensorflow object detection API is considered the best in practice for pre-training model within which Single Shot Detector MobileNet (SSD MobileNet v2 fpnlite 320x320 coco17 tpu 8) is downloaded and implemented. SSD Mobilenet is basically, an object detection model that computes the outlined box and classification of an object from a captured image. This Single Shot Detector (SSD) object detection model contains MobileNet as the base and followed by several convolutional layers as presented in fig.6. The functions of object localization and categorization are obtained in a single forward pass of the network. On the other hand, R-CNN series that obeys Regional Proposal Network (RPN) approach need two shots, one for the generation of region proposals and secondly, for identification of the object of each proposal. Therefore, SSD is assumed to be a lot faster as compared to two-shot RPN-based methods. This pre-trained model will allow us to perform the task called transfer learning a lot faster.

We have performed training the dataset up to 20,000

steps to obtain the best results. Finally, the trained model is tested in real time environment. The prediction of different signs such as hello, yes, bye, no, thank you is done using OpenCv and system camera. Table.1 represents the recognition rates of each static hand gesture when detected in real time.

### 3. Results and Discussion

S.No.	Labels assigned to hand signs	Recognition rate %
1.	HELLO	91.15
2.	YES	96.62
3.	NO	96.05
4.	THANK YOU	97.7
5.	BYE	92.47

Table 1 shows Labels and their recognition rates

Implementation of the proposed system was done on Jupyter notebook. From the Table. 1, the recognition rate of 'thank you' sign is the highest with 97.15%. While the prediction of 'hello ' sign shows the lowest recognition rate of 91.15 % because of its similarity to 'bye' hand gesture. 'bye' label was fluctuating with 'hello' due to high similarity in between the poses due to which its recognition rate is 92.47 %. The word label 'No' uses both hands for the gesture and obtained the recognition rate near to 96.05 %. The recognition rate of sign 'Yes' is 96.62%. The project has incurred satisfactory results and recognizes the word signs in static position. Overall performance rate of 94.798 % was achieved.

### 4. Conclusion

The system was successfully trained on all five words using transfer learning. This system will help us to detect five sign languages which are - Yes, Hello, Bye, No and Thank you. We obtained the average accuracy up to 94.798 % which is satisfactory. From this project, differently abled persons, especially deaf and dumb people will be able to communicate with the world effectively and at faster rate. Hence, it minimises the

communication barrier in between specialised people and common ones. The project can be further extended to detect different sign languages like ISL and ASL etc, number system and many other complete words, merely by adding their respective datasets and training down the models. The main objective of preparing a system that uses self-made hand gestures for different words, being used on daily basis, is successfully achieved.

## References

- [1] Vinothkumar, Krishnan Thangaiyan, Jayasankar. (2019). A Glove Based Approach to Recognize Indian Sign Languages.
- [2] K. Kudrisko, E. Flavin, X. Zhu and Q. Li, "Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review," in *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 82-97, 2021, doi: 10.1109/RBME.2020.3019769
- [3] Gupta, Himanshu Ramjiwal, Aniruddh Jose, Jasmin. (2018). Vision Based Approach to Sign Language Recognition. *International Journal of Advances in Applied Sciences*. 7. 156. 10.11591/ijaas.v7.i2.pp156161.
- [4] Tongi, Roman. (2021). Application of Transfer Learning to Sign Language Recognition using an Inflated 3D Deep Convolutional Neural Network.
- [5] Ismail, A.P. Aziz, Farah Kasim, NazirahDaud, Kamarulazhar. (2021). Hand gesture recognition on python and opencv. *IOP Conference Series: Materials Science and Engineering*. 1045. 012043. 10.1088/1757899X/1045/1/012043.
- [6] Fei Wang, Ronglin Hu, Ying Jin, Research on gesture image recognition method based on transfer learning, *Procedia Computer Science*, Volume 187, 2021,
- [7] Tolentino, Lean KarloSerfa Juan, Ronnie Thio-ac, August Pamahoy, Maria Forteza, Joni Garcia, Xavier. (2019). Static Sign Language Recognition Using Deep Learning. *International Journal of Machine Learning and Computing*. 9. 821-827. 10.18178/ijmlc.2019.9.6.879.
- [8] C J, SruthiLijija, A.. (2019). Signet: A Deep Learning based Indian Sign Language Recognition System. 0596-0600. 10.1109/ICCSP.2019.8698006.
- [9] Wario, Ruth Nyaga, Casam. (2018). Sign Language Gesture Recognition through Computer Vision.
- [10] Y. -C. Chiu, C. -Y. Tsai, M. -D. Ruan, G. -Y. Shen and T. T. Lee, "Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems," 2020 International Conference on System Science and Engineering (ICSSE), 2020, pp. 1-5, doi: 10.1109/ICSSE50014.2020.9219319.