



# **iJRASET**

International Journal For Research in  
Applied Science and Engineering Technology



---

# **INTERNATIONAL JOURNAL FOR RESEARCH**

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 10    Issue: IV    Month of publication: April 2022**

**DOI: <https://doi.org/10.22214/ijraset.2022.41625>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Hand Sign Recognition for Medical Field Genesis

Dr. Uttara Gogate<sup>1</sup>, Vahaz Kazi<sup>2</sup>, Aakash Jadhav<sup>3</sup>, Tirthesh Bari<sup>4</sup>

<sup>1</sup>Associate Professor, Computer Engineering Department, SSJCOE, Maharashtra

<sup>2, 3, 4</sup>Student, Computer Engineering Department, SSJCOE, Maharashtra

**Abstract:** *Inability to talk is taken into account to be true incapacity. folks with this incapacity use totally different modes to speak with others, there are variety of strategies accessible for his or her communication one such common methodology of communication is language.*

*Developing language application for speech impaired folks is vital, as they'll be ready to communicate simply with even people who don't perceive language.*

*Our project aims at taking the fundamental step in bridging the communication gap between traditional folks, deaf and dumb folk's victimization language. the most focus of this work is to form a vision-based system to spot language gestures in real-time. Not solely that, however our project conjointly aims to focus on the audience or the users with very little or no data regarding any language.*

*The project can aim to be helpful for each the traditional yet because the person with speaking or hearing incapacity. Although, the scope of the project shall be on the far side the compatibility, still it's herewith tried to usher in the module or a locality of the appliance for the speech impaired person. Future scope of the project aims to develop the appliance to assist person with all ability.*

**Keywords:** *Hand Sign Recognition, American-Sign-Language, Yolo, Object Detection, Gesture to Speech, CNN, Machine Learning.*

## I. INTRODUCTION

Since ages, the civilization has noticed the chromosomal mutation in numerous completely different means. Let that be the mutation resulting in impairment in speech, hearing, or travel. This creates an enormous distinction within the society, like to be able to work with the traditional individuals, communicate with them, etc. analysis conducted shows that, an enormous portion of the worldwide population has the lack of speaking either partly or utterly. To tackle this drawback of communication gap, linguistic communication was fictitious. linguistic communication is that the most natural and communicative means for the speech impaired individuals.

During medical treatment patients are exposed to high amount of anesthesia, due to this reason most of the times patients are unable to explain what issues they are facing regarding their health. So, with the help of hand sign recognition patient can elaborate or communicate with doctor. Sign Languages vary supported location, with one amongst the foremost common ones being yankee linguistic communication (ASL) and British linguistic communication (BSL)[2]. People, United Nations agency don't seem to be deaf, ne'er try and learn the linguistic communication for interacting with the deaf individuals. This ends up in isolation of the deaf individuals. therefore, to cut back the communication gap, the analysis within the field of gesture to speech (G2S) system becomes additional vital. to attain the task of gesture recognition in video pictures, image process and machine learning techniques are involved [6]. In recent years several researchers targeted available gesture detection and developed several techniques and solutions. The various project provides associate degree economical application that solves the matter to some extent. The software package is associate degree application that helps in economical communication between the speech impaired individuals and traditional individuals. The project consists of a central repository containing yankee linguistic communication (ASL) interpretations offered in conjunction with associated details. These details embrace linguistic communication hand implications, their associated meanings and their uses. These details facilitate in manipulating the message that's required to be passed on. The project is associate degree application code that converts the message to be passed on from the hand language i.e., Gesture to Speech (G2S). Moreover, the applying may be developed to perform the vice-versa i.e., Speech to Gesture(S2G). the applying runs in real time therefore on build it additional user friendly and additional usable.

## II. LITERATURE SURVEY

Paper topic	Methodology	Review	Limitation
Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model <sup>[1]</sup>	YOLOv3 (You Only Look Once) Algorithm.	Lightweight model based on YOLOv3 and DarkNet-53 Gesture recognition without additional preprocessing, image filtering, and enhancement of images. High accuracy in real-time	Data set in YOLO format High computational power YOLO Lite
Sign Language to Speech Translation <sup>[2]</sup>	Convolutional Neural Network Text-to-Speech Translator	Gesture-to-Text and then Text-to-Speech approach Simple implementation	No real-time implementation done. Accuracy is less.
Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation <sup>[3]</sup>	3DCNN Algorithm MLP concept	Hand segmentation based on the openpose framework Optimizing the level of C3D architecture for HGR	Complex implementation No user interface No Real-time implementation done.
Real-time Hand Gesture Communication System in Hindi for Speech and Hearing Impaired <sup>[4]</sup>	Used web cam. 32 binary number combination used for recognition. YCbCr and Canny Edge modules used.	Accuracy is very high. Flexible modification according to user needs.	Made using MATLAB. No use of AI / ML. Not universal language.
Research on the Hand Gesture Recognition Based on Deep Learning <sup>[5]</sup>	Camshift Algorithm used for gesture tracking. LeNet-5 network to recognize hand gestures. AdaBoost classifier used for strong learning algorithm.	Accuracy is medium up to 98%. Rotational hand gestures are also covered. Increased iteration decreases loss curve.	Unable to obtain 3d information. Requires proper background lightening
Real – time Two Hand Gesture Recognition with Condensation and Hidden Markov Models <sup>[6]</sup>	Gesture recognition system using Hidden Markov Models (HMM). AdaBoost classifier used for strong learning algorithm. Baum- Welch algorithm is used to training for initialized HMM	Provides good results for recognizing gesture in real-time. Multiple hand gestures can be recognized.	Requires lots of process to segment the required image. Currently not optimized for sign language detection. Only used for normal regular gestures
Combining Hand Detection and Gesture Recognition Algorithms for Minimizing Computational Cost <sup>[7]</sup>	Classical computer vision algorithms are used Skeleton model is used for static gesture’s recognizer combination of gesture recognition and hand detection.	it’s possible to efficiently use computing resources depending on the context Expected performance increase can be calculated in advance by the formula.	Processing time is increased significantly as compared to other previous models. Prediction accuracy is decreased slightly.
Messaging and Video Calling Application for Specially Abled people using Hand Gesture Recognition <sup>[8]</sup>	Dataset Preparation Convert color images to grayscale Region-based Convolutional Neural Network	Average Time of YOLO v4 is remarkably less the scope can be extended for blind people also.	In a vision-based approach CNN works better.
Real-Time Recognition of Sign Language Gestures and Air-Writing using Leap Motion <sup>[9]</sup>	Preprocessing and Feature Extraction using Normalization Initial Classification Using SVM BLSTM-NN classifier-based Gesture Recognition	An accuracy of 100% has been recorded using SVM classifier. An overall accuracy of 63.57% has been recorded	The accuracy of the system is low because it has been tested with a lexicon free approach.

Table I. Literature Survey

Hence, from the literature survey we can infer that for the project it would be a feasible option to convert the image to Grayscale from RGB. Along with that it can also be concluded that using CNN would help to increase the efficiency in real-time. Using any hardware to detect the gestures can make the system complex and not user-friendly. In order to make system more efficient and sustainable, it is feasible to use the hand gestures using normal convections.

### III. PROBLEM STATEMENT

As we know some people unfortunately, accidentally or by birth bear some vocal issues due to which communicating with normal people becomes a difficult problem for them. Hence, communicating with the normal person only takes place via hand gestures. As of today, one third of the world's population uses ASL i.e., American Sign Language while communicating with the speech impaired person. But not everyone is aware about the gestures of the alphabet in ASL. Due to which communicating becomes a difficult task for the common person. Not only that, but also it becomes relevantly difficult for the impaired person to make the normal person understand his point of view. This creates a hole in communication. Hence, we need to develop a solution for the problem being a communication gap, so as to make communication easier for both the impaired as well as normal people. Such solution shall help them to stand and be heard in the crowd. Hence, the problem statement is developed, to bridge the communication gap created. It aims to make the person with no understanding of ASL to be able to recognize/rectify the impaired person's point of view. It also aims to provide a system that is user friendly, and ease the communication, along with that works in real time.

### IV. PROPOSED SYSTEM

Based on the higher than survey we've got shortlisted to figure on YOLO Machine learning algorithmic rule. There are unit totally different versions of yolo model accessible to figure with and in keeping with the survey, system can sway work with efficiency with YOLO Version three.0. YOLO could be a Convolutional Neural Network (CNN) for acting object detection in period. CNNs area unit classifier-based systems that may method input pictures as structured arrays of information and establish patterns between them. YOLO has the advantage of being abundant quicker than alternative networks and still maintains accuracy. It permits the model to appear at the total image at check time, thus its predictions area unit hip by the worldwide context within the image. YOLO and alternative convolutional neural network algorithms "score" regions supported their similarities to predefined categories. The planned system can have twenty-six categories for every alphabet within the English. The dataset needed for the YOLO-V3 implementation is meant to be in YOLO format. Hence, the pictures that area unit used for coaching and testing of the model.

The most salient feature of v3 is that it makes detections at 3 totally different scales. YOLO may be a totally convolutional network and its ultimate output is generated by applying a one x one kernel on a feature map. In YOLO v3, the detection is completed by applying one x one detection kernels on feature maps of 3 {different|totally totally different|completely different} sizes at 3 different places within the network. The shape of the detection kernel is one x 1 x (B x (5 + C)). Here B is that the variety of bounding boxes a cell on the feature map will predict, "5" is for the four bounding box attributes and one object confidence, and C is that the variety of categories. In YOLO v3 trained on coconut palm, B = three and C = eighty, therefore the kernel size is one x one x 255. The feature map created by this kernel has identical height and breadth of the previous feature map, and has detection attributes on the depth as represented on top of. YOLO v3 makes prediction at 3 scales, that area unit exactly given by down sampling the scale of the input image by thirty-two, sixteen and eight severally. The 82nd layer created the first detection. For the primary eighty-one layers, the image is down sampled by the network, specified the 81st layer contains a stride of thirty-two. If we've a picture of 416 x 416, the resultant feature map would be of size thirteen x thirteen. One detection is created here mistreatment the one x one detection kernel, giving North American country a detection feature map of thirteen x thirteen x 255. Then, the feature map from layer seventy-nine is subjected to some convolutional layers before being up sampled by 2x to dimensions of twenty-six x twenty-six. This feature map is then depth concatenated with the feature map from layer sixty-one. Then the combined feature maps are once more subjecte3d someone x one convolutional layers to fuse the options from the sooner layer (61). Then, the second detection is created by the 94th layer, yielding a detection feature map of twenty-six x twenty-six x 255. A similar procedure is followed once more, wherever the feature map from layer ninety-one is subjected to few convolutional layers before being depth concatenated with a feature map from layer thirty-six. Like before, someone x one convolutional layers follow to fuse the data from the previous layer (36). we tend to build the ultimate of the three at 106th layer, yielding feature map of size fifty-two x fifty-two x 255. YOLO v3 area unit currently expected through provision regression. YOLO v3 performs at par with different state of art detectors like RetinaNet, whereas being significantly quicker, at COCO mAP fifty benchmark

### V. CONCLUSION

This application can be developed using the CNN model successfully and was tested successfully by taking few test cases. It is user friendly and has required options, which can be used by user to perform the speech conversion via gesture. According to the result analysis the current model works properly and has achieved the goal of getting 95% accuracy within the desired format. We have implemented various best practices to create and train our model.

Throughout the development of the model, we have learned various best practices and architecture patterns being used in industry today. Using YOLOV3 will prove to be more efficient in terms of results and accuracy. Optimum utilization of resources will be done. Efficient management of records can be achieved. Simplification of Operations as the process follows simple convention and flow. Less processing time and ease of getting required information. Usefulness and correctness of the system provides better user experience.

### REFERENCES

- [1] Abdullah Mujahid, Mazhar Javed Awan, Awais Yasin, Mazin Abed Mohammed, Robertas Damaševičius, Rytis Maskeliunas and Karrar Hameed Abdulkareem, Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model, Appl. Sci. 2021.
- [2] Aishwarya Sharma, Dr. Siba Panda, Prof. Saurav Verma, Sign Language to Speech Translation, ICCCNT 2020.
- [3] MUNEER AL-HAMMADI, GHULAM MUHAMMAD, WADOOD ABDUL, MANSOUR ALSULAIMAN, MOHAMMED A. BENCHERIF, TAREQ S. ALRAYES, HASSAN MATHKOU, AND MOHAMED AMINE MEKHTICHE, Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation, IEEE Access 2020
- [4] Shilpa Chaman, Dylan D'souza, Benz D'mello, Karan Bhavsar, Jolton D'souza, Real-time Hand Gesture Communication System in Hindi for Speech and Hearing Impaired, ICICCS 2018.
- [5] Jing-Hao Sun, Jia-Kui Yang, Ting-Ting Ji, Guang-Rong Ji, Shu-Bin Zhang, Research on the Hand Gesture Recognition Based on Deep Learning, IEEE 2018.
- [6] Tanatcha Chaikhumphapha, Phattanaphong Chomphuwiset, Real – time Two Hand Gesture Recognition with Condensation and Hidden Markov Models, IEEE 2018.
- [7] Roman Golovanov, Dmitry Vorotnev, Darina Kalina, Combining Hand Detection and Gesture Recognition Algorithms for Minimizing Computational Cost, IEEE XPLORE 2020.
- [8] Rachana R. Chhajed, Komal P. Parma, Manvi D. Pandya, Neha G. Jaju, Messaging and Video Calling Application for Specially Abled people using Hand Gesture Recognition, I2CT 2021
- [9] Pradeep Kumar and Rajkumar Saini, Santosh Kumar Behera and Debi Prosad Dogra, Partha Pratim Roy, Real-Time Recognition of Sign Language Gestures and Air-Writing using Leap Motion, IAPR CONFERENCE 2017



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)