iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42961

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Real-Time Sign Language Detection

Subhashini Yadav, Shreyashi Raj, Kashish Awasthi, Rohit Bidwan, Lokesh Jain
Inderprastha Engineering College Ghaziabad

Abstract: Communication is the method of sharing or exchanging information, ideas or feelings. To have a communication between two people, both of them need to have knowledge and understanding of a common language. But in the case of deaf and dumb people, the means they use for communicating is different from that of normal people. Deaf is not able to hear and dumb is not able to speak.

They communicate using sign language among themselves and with normal people but normal people don't take seriously the importance of sign languages. Not every-one has the knowledge and understanding of sign language which make the communication difficult between a normal person and a deaf and dumb person. For overcoming this barrier,a model can be build based on machine learning.

A model can be trained to recognize different gestures of sign language and translate them into English language. This will help a lot of people in communicating with deaf and dumb people with ease. A real time ML based system is built for the real time sign language detection with TensorFlow object detection in this paper. The major purpose of this project is to build a system for the differently abled people to communicate with others easily and efficiently.

Keywords: Object detection, sign language, Deep learning, SSD(Single-shot Detector), TensorFlow

I. INTRODUCTION

Communication can be defined as the process of transferring information from one place, person, or group to another place,person or group. It consists of three components: the speaker, the message that is to be communicated, and the listener. Communication can be considered successful only when whatever message the speaker is trying to convey is received and understood by the listener.

In this paper a real time ML based system is built for the real time sign language detection with TensorFlow object detection in this paper. This model is using SSD ML algorithm, by recognizing the signs as words instead of old traditional translators, that are very slow and take too much since every alphabet has to be recognized to form the whole sentence in old methods.

TensorFlow object detection API is a powerful library that can enable anyone to build and deploy powerful image recognition system. This object detection includes recognizing the objects and classifying the objects and then localizing those objects and attract the bounding boxes surrounding them.

TensorFlow has a broad range of attention and use in the field of machine learning globally, because the second generation is learning the system of google. TensorFlow has the advantage of elasticity and high attainability . Deep learning is the base of the object detection algorithm which is also very convenient to execute through TensorFlow, and the hardware environment needs are reasonable, which is very suitable for the research of this paper.

This research Paper focuses on a real time ML based system that is built for the real time sign language detection with TensorFlow object detection. This model is using SSD ML algorithm, by recognizing the signs as words instead of old traditional translators, that are very slow and take too much since every alphabet has to be recognized to form the whole sentence in old methods.

II. LITERATURE SURVEY

Lots and lots of work has been happened in this area from years and all of it had continuous growth and improvements in this object detection field. So, we have first written about the different applications of detection of objects and then the implementation algorithm of those systems or projects.

1) The first ever approach to sign language detection was done by Bergh in 2011 . A hand gesture recognition system was built which gave good results but it only considered six gestures.
2) Balbin et al. recognized that the system was only recognizing five Filipino words and was using colored hand gloves for hand position recognition. Our model is trained to detect different hand gestures without any colored gloves, using bare hands only.
3) Apoorva Raghunandan , did the work on different algorithms like skin, color and face detection, they simulated and implemented different objects with the MATLAB in video surveillance applications. They used Viola jones algorithm for the face detection in their model. It detected all the features of the face like eyes, nose, ears, etc. They gave the accuracy of 95% with various object detection algorithms and simulated that in the MATLAB.

©IJRASET: All Rights are Reserved

4) Krizhevsky released AlexaNet which used deep learning in various aspects for object detection in 2012.

5) Girshick et al. showed the advantages of neural networks for detecting objects naming R- CNN. But, it needed huge datasets and the datasets available at that time were very low. They used ImageNet 2012 dataset to pre-train the system which solved the scarcity problem of dataset. Garshick proposed a faster object detection algorithm in 2015 called Fast R-CNN in which, image is inputted first to a single CNN with many convolutional layers that generates a convolution feature map. The advantage of this was that it trained the entire image with only one CNN other than training images with multiple CNNs for all the region of the image.

6) The SSD model was first adopted for hand detection with proposed model based IsoGD dataset, which gave the accuracy of 4.25% . Real-time object detection came in view in 2016 in testing images and two algorithms were proposed naming YOLO and SSD. YOLO used CNN to reduce spatial dimension detection box and performed linear regression and made boundary box predictions. While in SSD, the size of the box that detects is usually fixed and used for simultaneous size detection. SSD advantage is simultaneous detection of objects with different sizes.

III. PROPOSED METHODOLOGY

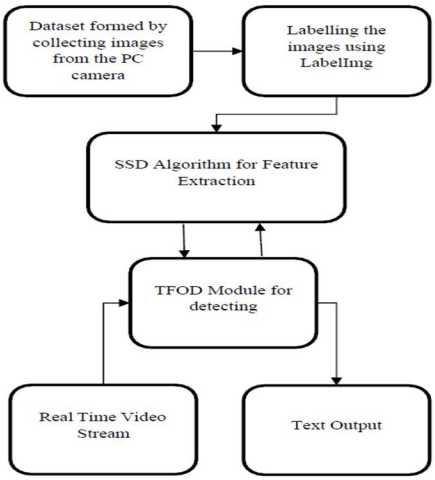
The proposed system has been designed to develop a real-time sign language detector using a TensorFlow object detection API and train it through deep learning for the created dataset . For dataset creation, images have been captured by webcam of laptop using Python and OpenCV . After the data acquisition, a labeled map is created which is a representation of all the objects within the model, i.e., it contains the label of each sign along with their ids. These will be used as a reference to look up for the class name. TF records of the training data and the testing data are then created using generate_tfrecord which is used to train the TensorFlow object detection API.Binary storage format of TensorFlow is TF record. Usage of binary files for storage of the data significantly impacts the performance of the import pipeline and also, the training time of the model. It takes less space on the disk, copies fastly and easily, and can be read efficiently from the disk. The open-source framework, TensorFlow object detection API makes it easy to develop, train and deploy an object detection model. They have their framework called the TensorFlow detection model zoo which offers various models for detection that have been pre-trained on the COCO 2017 dataset.

This open-source framework, TensorFlow object detection API makes it easier to develop, train and deploy an object detection model. TensorFlow object detection API have their framework called the TensorFlow detection model zoo that offers various models for detection that have been pre-trained or already trained on the COCO 2017 dataset. The pre-trained TensorFlow model that is used here is SSD MobileNet v2 320x320.

The real-time detection is being done using OpenCV and webcam of the laptop again. For, real-time detection, cv2, and NumPy libraries of python are used. The system detects signs in real-time and translates what each of the gesture means into English language . The system

Has been tested in real-time by creating and showing it different signs. The confidence rate of

each sign , i.e., how confident the system is in recognizing a sign is checked, noted, and tabulated for the result.



System Architecture

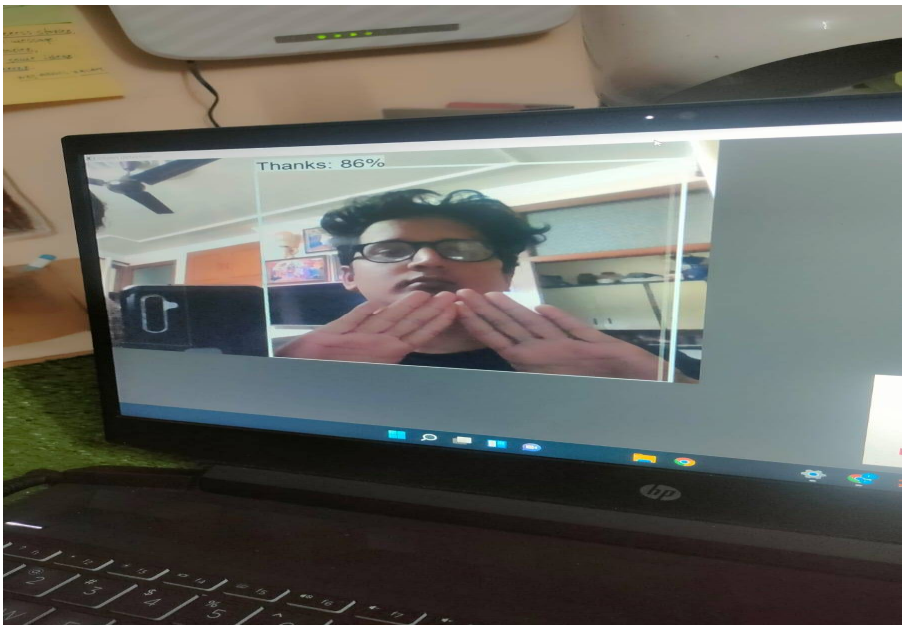
For data acquisition, images have been captured by webcam of laptop using Python and OpenCV. OpenCV has functions which are primarily aimed at the real-time computer vision. It escalates the use of machine perception in the commercial products and provides a common infrastructure for the applications based on computer visions. The OpenCV library has more than 2400 efficient computer vision and machine learning algorithms. These algorithms could be used for face detections and recognitions, object identification, classification of human actions, camera tracking and object movements,and many more. Once all of the images have been captured by the webcam, they are then one by one labelled using the LabelImg software. LabelImg is a free open-source tool that is for graphically labelling images. When the labelled image is saved, its XML file is created. These XML files have all the details of the images and also the details of the labelled portion. Once labelling of all the images is done, their XML files are available. This is used for creating the TensorFlow records(TF records). All the images and their XML files are then divided into training data and validation data in the ratio of 80:20.

The data samples are collected for 4 words and one sentence . The data samples are recorded by us, using a digital camera.

1) Hello
2) Thanks
3) I Love you
4) How
5) No.

IV. RESULT AND CONCLUSION

Real coloring pictures from a PC camera were used in a real-time sign language detection with a SSD algorithm was presented. In this paper, to help differently able people in easy interfacing with others, signs are translated into text statement. This took advantage of deep learning techniques and showed good results. The outcomes obtained by the system are discussed in this system.
Sign languages are kinds of visual languages that employ movements of hands, body, Sign languages are very important for specially-abled people to have a way of communication. Through sign languages, specially-abled can communicate ,express and can share their feelings with others. The drawback is that not everyone has the knowledge of sign languages which limits communication between specially-abled and normal people. This limitation could be overcome with the use of automated Sign Language Recognition systems that will be able to easily translate the sign language gestures into commonly spoken languages such as English language. In this paper, it has been done by using the TensorFlow object detection API. The system detects sign language in real-time. For data acquisition, images have been captured by webcam of laptop using Python and OpenCV which makes the cost cheaper. The developed system is showing an average confidence rate of 80.25%. Though the system has achieved a high average confidence rate, the dataset on which it has been trained on is small in size.



V. FUTURE SCOPE

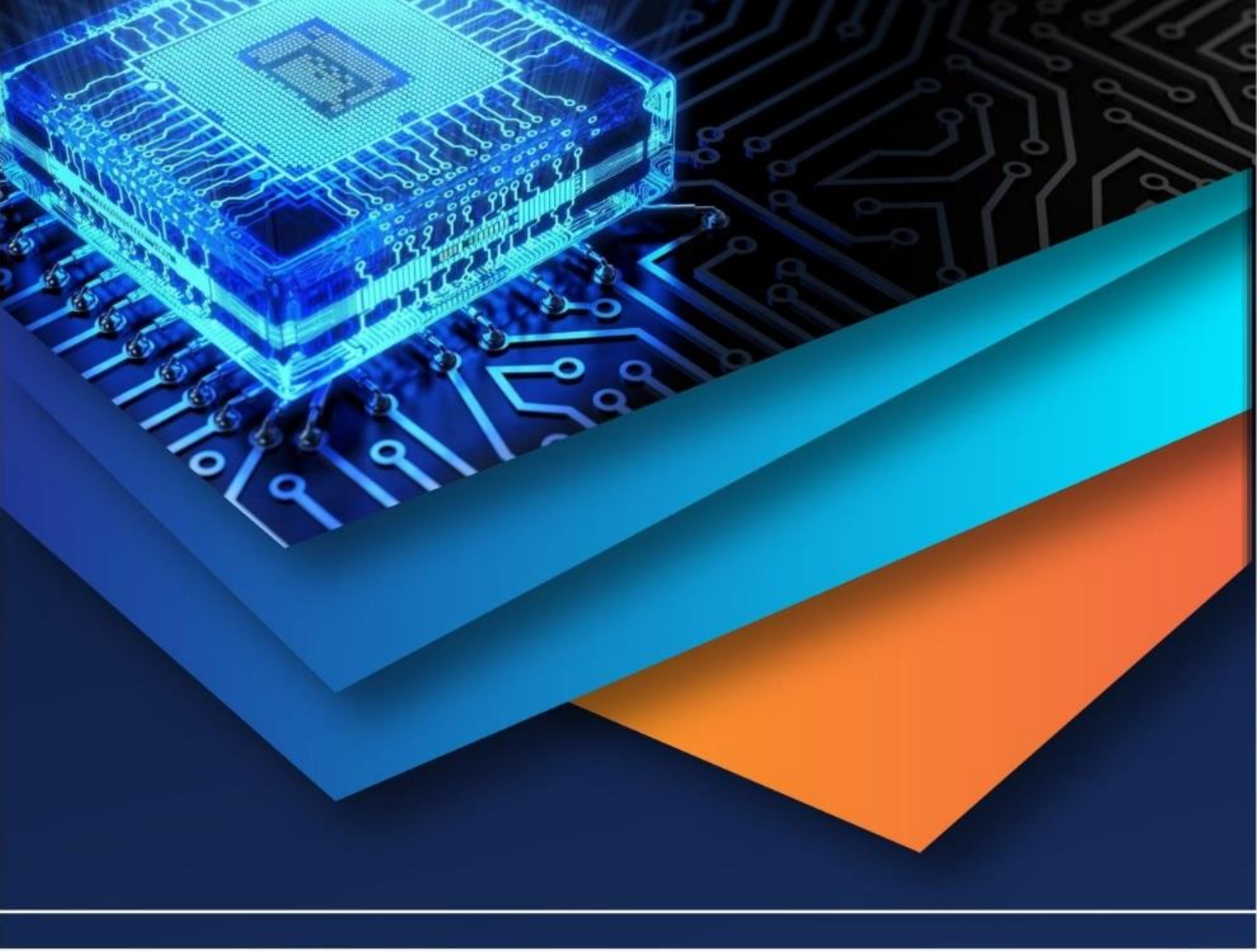
In future, the dataset that has been used can be enlarged so that the system can recognize more gestures. The model which is Tensorflow model that has been used can be replaced with another model as well. The same system can be implemented for different sign languages by substituting the dataset.

VI. ACKNOWLEDGEMENT

We express our deepest gratitude and heartfelt thanks to our mentor, DR.LOKESH JAIN(Information Technology Department), for his expert guidance, constant encouragement, constructive criticism, and inspiring advice throughout the completion of this report.

REFERENCES

[1] M. Van den Bergh and L. Van Gool, "Combining RGB and ToF cameras for real-time 3D hand gesture interaction," 2011 IEEE Workshop on Applications of Computer Vision (WACV), 2011, pp. 66-72
[2] J. R. Balbin et al., "Sign language word translator using Neural Networks for the Aurally Impaired as a tool for communication," 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2016, pp. 425-429
[3] Raghunandan, Apoorva, Pakala Raghav, and HV Ravish Aradhya. "Object Detection Algorithms for video surveillance applications." In 2018 International Conference on Communication and Signal Processing (ICCSP), pp. 0563- 0568. IEEE, 2018
[4] Alex Krizhevsky, Ilya Sutskever, and Geo_rey Hinton. Imagenet classi_cation with deep convolutional neural networks. In Advances in Neural Information Processing Systems, pages 1097-1105, 2012.
[5] Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587. 2014
[6] Wu W, Dasgupta D, Ramirez E, et al. Classification accuracies of physical activities using smartphone motion sensors[J]. Journal of medical Internet research,2012,14(5): el30-el30
[7] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. Curran Associates Inc.2012:1097-1105.
[8] D. Mart, Sign Language Translator Using Microsoft Kinect XBOX 360 TM, 2012, pp. 1-76.



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)