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.2022.41924

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 IV Apr 2022- Available at www.ijraset.com

American Sign Language Translator

Mohammed Nihad K[1], Sarath Suresh C[2], Anirudhan SS[3]
[1, 2, 3]*B.Tech Computer Science & Engineering, Lovely Professional University, Phagwara, Punjab, India*

Abstract: Sign language is one of the ancient and most natural form of language for communication. Sign Language is being used by many people in the world, but only a small part of the population knows how to interpret the language. Which makes it tough for the Deaf people to communicate with other people. We used the Tensorflow object detection API in our research, which is an open-source framework for training and deploying object detection models.
Keywords: TensorFlow Object Detection API

I. INTRODUCTION

American Sign Language (ASL) is a natural language that serves as the predominant sign language of Deaf communities around the world. ASL has the same language properties as spoken languages, with syntax that is different from English. ASL is 'spoken' by actions of the hands and facial expressions.

In our project we basically focused on creating a model that can recognise hand gestures, which will be detected by the program as input via a Webcam which captures the human gestures with the help of OpenCV. Then show us the corresponding word we already labelled to that gesture as output. Image labelling is done with the help of LabelImg, a graphical Annotation Tool.

Dumb and Deaf people mostly rely on sign language interpreters for communications. This model reads your hand gestures and translate them to English words using Tensorflow Object Detetction API.

In this model , we trained 5 gestures, but we always train more with this prototype model. The 6 gestures are as follows:

1) Hello
2) I Love You
3) Thank you
4) Yes ✔
5) No ☒

II. MOTIVATION

There is always a language barrier between deaf and dumb people and other people who can communicate verbally. So they rely on sign languages for basic communication. If only there was a common interface where gestures could be easily understood by other people, communication would be much easier. So the research has always been done for a gesture recognition interface where D&M people can enjoy communicating without the language barrier. The idea is to develop user-friendly Human Computer Interfaces (HCI) where the computer understands American Sign Language.

III. KEY WORDS AND DEFINITIONS

A. TensorFlow

Tensorflow is a numerical computing software package that is open source. We define the nodes of the computation graph first, and then the actual computation takes place within a session. TensorFlow is a popular library.

B. Keras

Keras is a Python toolkit for high-level neural networks that acts as a wrapper for TensorFlow. It's useful when we need to build and test a neural network rapidly and with few lines of code. It includes implementations of typical neural network parts such as layers, objectives, activation functions, optimizers, and tools for working with images and text data.

C. OpenCV

OpenCV (Open-Source Computer Vision) is a programming library for real-time computer vision that is open source. Image processing, video recording, and analysis for features such as face and object identification are the most common applications. It is written in C++ and uses that language as its primary interface.

D. LabelImg

LabelImg is a tool for labelling graphical images.

It's written in Python and has a Qt-based user interface. Annotations are saved as a XML file in the PASCAL VOC format, the format used by ImageNet. In addition, it also supports YOLO and CreateML formats.

IV. METHODOLOGY

The suggested system is intended to generate a real-time sign language detector using a TensorFlow object detection API and train it for the created dataset using transfer learning. Images are taken using a camera with Python and OpenCV for data collection.

V. DATA ACQUISITION

For American Sign Language, a real-time sign language detecting system is being developed. Images are taken using a camera with Python and OpenCV for data collection. OpenCV has a lot of capabilities that are geared toward real-time computer vision. It makes machine perception more accessible in commercial goods and provides a common architecture for computer vision-based applications. Over 2500 efficient computer vision and machine learning algorithms are available in the OpenCV library, which can be used for face detection and recognition, object identification, classification of human actions, tracking camera and object movements, extracting 3D object models, and many other tasks.

The dataset is made up of 15 photos for each gesture.The photos are collected every 2 seconds, allowing time to record gestures with slight variations each time, and a five-second pause is supplied between two individual signs, i.e., a five-second interval is provided to shift the sign of one alphabet to the sign of a different alphabet. The photographs are saved in the appropriate folder.

Dependencies such as cv2, i.e., OpenCV, OS, time, and uuid have been imported for data acquisition. To work with file paths, the dependency OS is used. It is included in the Python standard utility modules and provides functions for dealing with operating systems. Time can be represented in Python code in a variety of ways, including objects, numbers, and texts, thanks to the time module. It can be used to quantify code efficiency or wait during code execution, in addition to representing time. It's being used here to insert gaps between image captures to allow for hand motions. The image files are named using the uuid library. It aids in the production of 128-bit random objects as ids are produced based on time and computer hardware, they provide uniqueness.

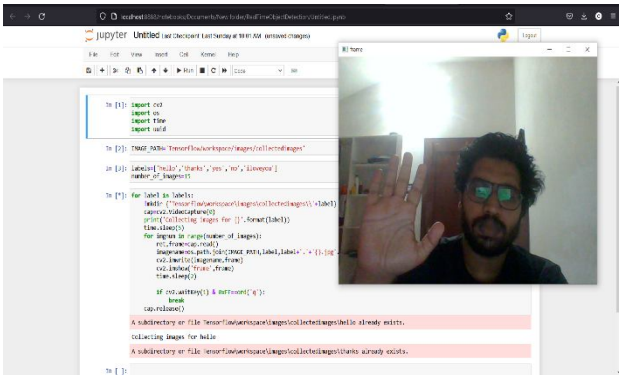


Figure.1 Collecting image for ‘Hello’ using OpenCV

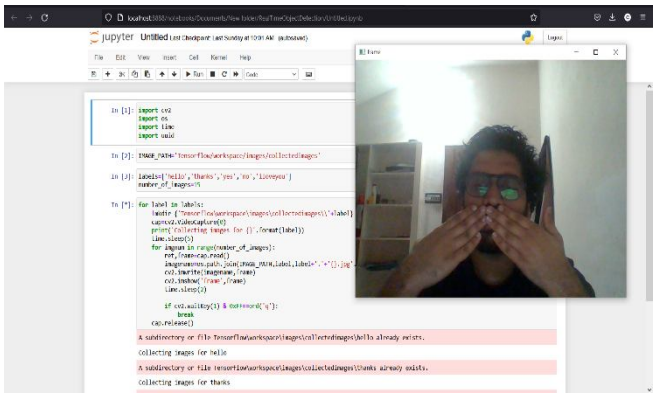


Figure 2. Collecting images for ‘Thanks’ using OpenCV

Similarly, we collect images for all 5 gestures using OpenCV and save them in separate folders.

After all of the photos have been acquired, the LabelImg package is used to label them one by one. LabelImg is a free open-source programme for labelling images graphically. The XML file for the labelled image is created when it is saved. The XML files contain all of the image details, including the labelled portion's details. After you've labelled all of the photographs, you'll have access to their XML files. This is where the TF (TensorFlow) records are created. After that, all of the photos, together with their XML files, are separated into training and validation data in an 80:20 ratio. Ten (80%) of the 15 photos of a gesture were taken and stored as a training dataset, while the other five (20%) were taken and stored as a validation dataset. This job was completed for each of the five gesture photos.

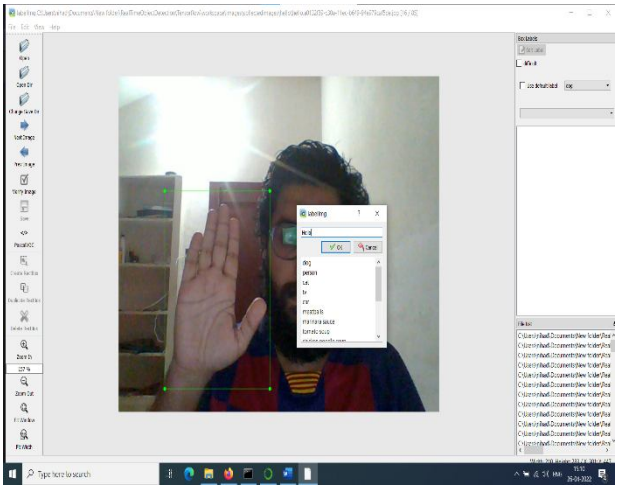


Figure 3. Annotating the image for ‘Hello’ using LabelImg

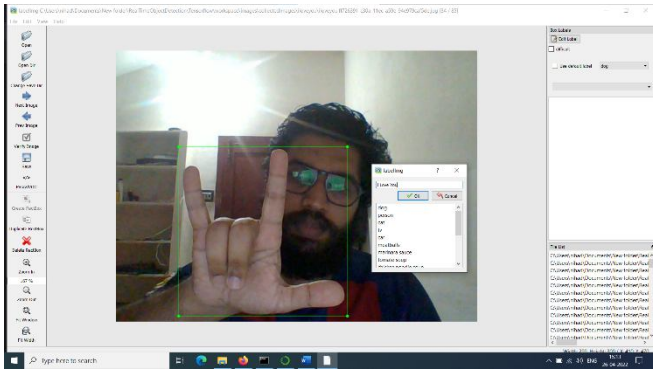


Figure 4. Annotating the image for ‘I Love You’ using LabelImg

VI. TESTING AND TRAINING

Following the data acquisition, a labelled map is constructed as a representation of all the objects in the model, containing the label of each sign as well as their id. The alphabet is represented by each of the five labels on the label map. A unique id ranging from 1 to 5 has been assigned to each label. This will be used to look up the class name as a reference. Then, using generate tfrecord, TF records containing the training and testing data are created, which are used to train the TensorFlow object detection API. TensorFlow's binary storage format is called TF record. As a result, using binary files for data storage has a substantial impact on the performance of the import process.

TensorFlow object detection API is an open-source platform that enables developing, training, and deploying object detection models simple. They provide a framework called the TensorFlow detection model zoo that includes a variety of detection models that have been pre-trained on the COCO 2017 dataset. SSD MobileNet v2 320x320 is the pre-trained TensorFlow model that is being used.

The live detection is done using OpenCV and webcam. For, real-time detection, cv2, and NumPy dependencies are used. The system detects signs in real time and translates what each gesture means into English. The system is tested in real-time by creating and showing it different signs.

VII.EXPERIMENTAL EVALUATION

The dataset is created for American Sign Language where signs are words of the English language. The dataset is created following the data acquisition method. The experimentation was carried out on a system with an Intel i5 8 th generation 2.70 GHz processor, 4 GB memory and webcam (HP TrueVision HD camera), running Windows 10 operating system. The programming environment includes Python (version 3.7.1), Jupyter Notebook, OpenCV (version 4.2.0), TensorFlow Object Detection API.

VIII. FINAL RESULT

In real time, the created system can detect American Sign Language and convert to text. TensorFlow object detection API was used to build the system. SSD MobileNet v2 320x320 is the pre-trained model that was pulled from the TensorFlow model zoo. On the produced dataset, which has 75 photos in total, 15 images for each gesture, it was trained using transfer learning.

IX. CONCLUSION

Sign languages are visual languages that communicate via the use of hand, body, and face expressive motions. Specially abled persons require a means of communication, which sign languages provide. They may communicate, express, and share their feelings with others through it. The disadvantage is that not everyone is familiar with sign languages, which limits communication. This constraint can be solved by employing automatic Sign Language Recognition systems that can easily translate sign English motions into spoken language. The TensorFlow object detection API was used in this paper. The American Sign Language alphabet dataset was used to train the system. In real time, the system recognizes sign language. For data acquisition.

REFERENCES

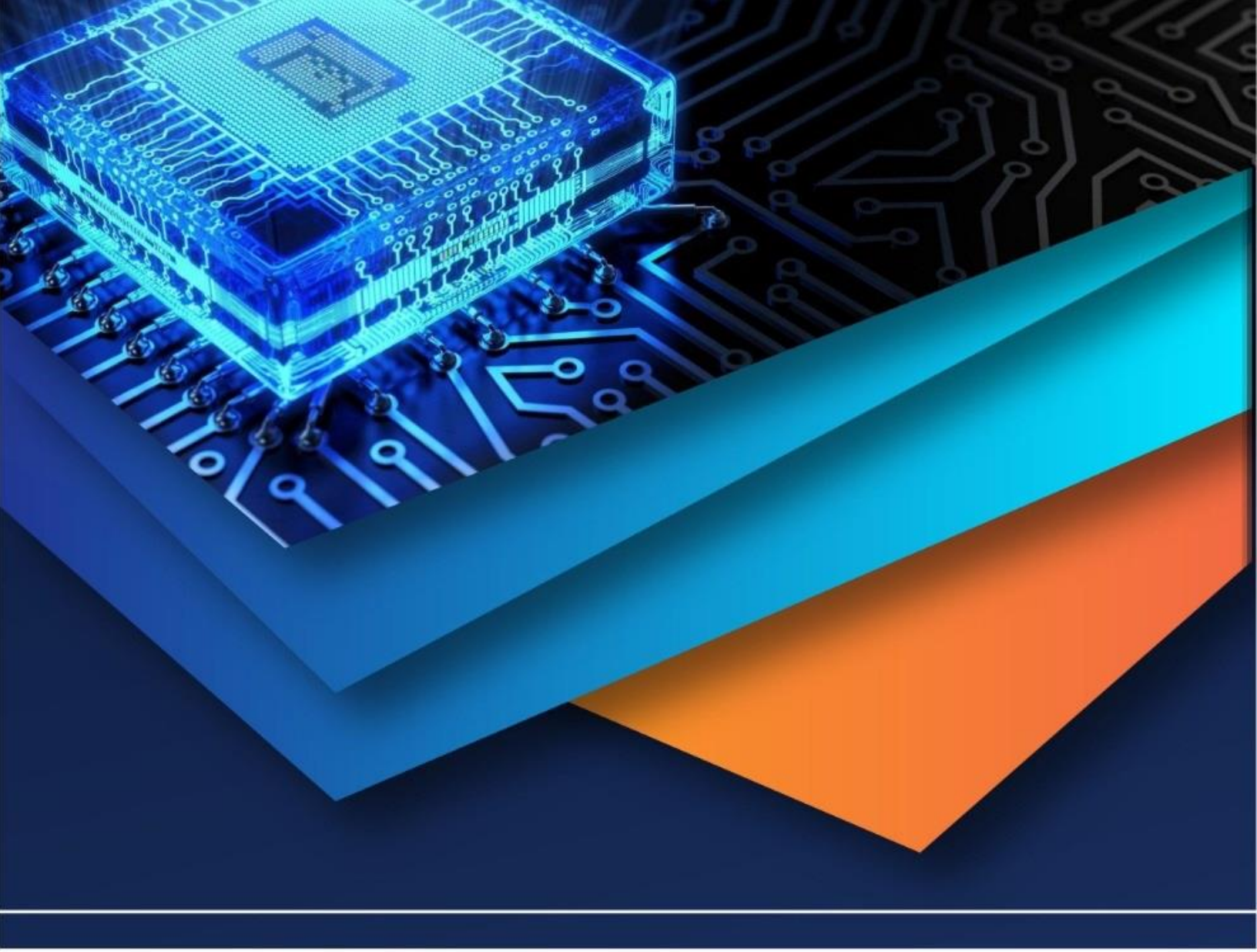
[1] Suharjito, Anderson, R., Wiryana, F.,Ariesta, M.C., Kusuma, G.P.: Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input- Process-Output. Procedia Comput. Sci.116,441–448 (2017). https://doi.org/10.1016/J.PROCS.2017.10.028.

[2] Hu Peng, “Application Research on Face Detection Technology based on Open CV in Mobile Augmented Reality”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 2 (2015).

[3] Prof. Rajeshri Rahul Itkarkar, “A Study of Vision Based Hand Gesture Recognition for Human Machine Interaction”, International Journal of Innovative Research in Advanced Engineering, Vol. 1, pp:12, (2014).

[4] Ruize Xu, Shengli Zhou, Li, W.J, “MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition”, IEEE Sensors Journal, vol.12, no.5, pp.1166-1173, (2012).

[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



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY