IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

Sign Language Recognition using OpenCV

Amey Mete[1], Ambraj Yeldandi[2], Harshal Sable[3], Muskan Gavasekar[4], Prof. Sujay Pawar[5]
Students, Department of Information Technology[1,2,3,4]
Assistant Professor, Department of Information Technology[5]
Dr. D. Y. Patil Institute of Technology, Pune, Maharashtra, India

Abstract: *This paper focuses on experimenting with different segmentation approaches and unsupervised learning algorithms to create an accurate sign language recognition model. To more easily approach the problem and obtain reasonable results, we experimented with just up to 10 different classes in the our self-made dataset instead of all 26 possible letters. We collected more than 20 RGB images for each hand gesture and their corresponding depth data using SSD Mobile Net Model. Hand gesture is one of the method used in sign language for non-verbal communication. Various sign language systems has been developed by many makers around the world but they are neither flexible nor cost-effective for the end users.*

Keywords: Computer Vision, Object Detection, Desktop Application, Dataset

I. INTRODUCTION

The goal of this project was to build a network able to classify which signor gesture is being signed, given an image of a signing hand. This project is a first step towards building a possible sign language translator, which can take communications in sign language and translate them into written or text format. Such a translator would greatly lower the barrier for many deaf and mute individuals to be able to better communicate with others in day to day interactions. This goal is further motivated by the isolation that is felt within the deaf community. Loneliness and depression exists in higher rates among the deaf population, especially when they are immersed in a hearing world. Large barriers that profoundly affect life quality stem from the communication disconnect between the deaf and the hearing. Some examples are information deprivation, limitation of social connections, and difficulty integrating in society.

Most research implementations for this task have used depth maps generated by depth camera and high resolution images. The objective of this project was to see if networks are able to classify sign letters using simple images of hands taken with a personal device such as a laptop webcam. This is in alignment with the motivation as this would make a future implementation of a real time written language translator practical in an everyday situation.

- The sign language is a very important way of communication for deaf-dumb people. In sign language each gesture has a specific meaning. So therefore complex meanings can be explain by the help of combination of various basic elements
- Typically Sign recognition is related as image understanding. So here we are going to use Tensorflow Object Detection API Model using Python.
- The system will use a webcam for the capturing images and pre-processing of the signs will be done by using Jupyter Notebook as an IDE and OpenCv library.
- Using SMD MobileNet Model to recognise and train the model . it is a Single Shot Detector which is a pre-defined algorithm given by Tensorflow Object Detection API

The domain analysis that we have done for the project mainly involved understanding the SMD MobileNet Model. TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google. TensorFlow provides stable Python and C APIs and without API backwards compatibility guarantee: C++, Go, Java, JavaScript and Swift. Third-party packages are available for C#, Haskell Julia, MATLAB,R, Scala, Rust, OCaml, and Crystal."New language support should be built on top of the C API. However, not all functionality is available in C yet.

Some more functionality is provided by the Python API. Among the applications for which TensorFlow is the foundation, are automated image-captioning software, zoo. OpenCV (Open Source Computer Vision Library) is a library of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

programming functions mainly aimed at real-time computer vision and data science as well. Originally developed by Intel, it was later supported by Willow Garage then Itseez. The library is cross-platform and free for use under the open-source BSD license. OpenCV's application areas includes:

- 2D and 3D feature toolkits
- Egomotion estimation
- Facial recognition system
- Gesture recognition
- Human–computer interaction (HCI)
- Mobile robotics
- Motion understanding
- Object identification
- Segmentation and recognition

II. METHODOLOGY

We tried implementing many different versions of quick image segmentation, in particular, hand segmentation. Edge Segmentation One of the methods we tried was using a Canny edge detector to find relevant ”objects” in the field of view of the camera. The edges were then dilated, and then all remaining holes in the mask were filled to create a solid, continuous mask. Once this was done, only the largest areas were taken in order to remove all the background clutter objects. This approach makes the simplifying assumption that the biggest objects seen in segmentation are typically of the most interest as well. Sign language recognition is an important application of gesture recognition. Sign language recognition has two different approaches.

2.1 Glove Based Approach

- In data Glove based method the sensors in the glove can detect the movement of hands and pass this information to the computer. This approach has high accuracy in gesture recognition but it is quite expensive and inconvenient to the user.

2.2 Vision Based Approach

- Vision based methods provide more user convenience. It will give better accuracy at higher computation rate. Vision based methods involve various image processing and pattern classification techniques for sign language recognition.
- Due to the limitations of glove-based, most of the research works focus on appearance-based approach

III. LITERATURE REVIEW

In gesture recognition system various technologies are used for capturing hand gestures, which includes a Glove-based method and Vision based method. This method uses Convolutional Neural Networks (CNN) to generate a low dimensional representation called embeddings. Then those embeddings are used to classify the person’s facial images. By this system different types of applications like student attendance system, security etc can be developed. In data Glove based method the sensors in the glove can detect the movement of hands and pass this information to the computer. This approach has high accuracy in gesture recognition but it is quite expensive and inconvenient to the user. Vision based methods provide more user convenience. It will give better accuracy at higher computation rate. Vision based methods involve various image processing and pattern classification techniques for sign language recognition.

IV. WORKING OF PROPOSED SYSTEM

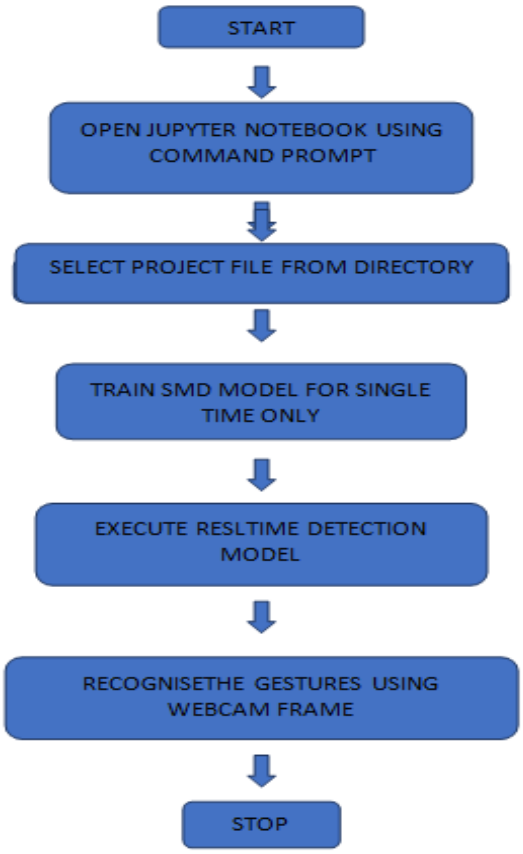
In this project, we have implemented an automatic sign language gesture recognition system in real-time, using tools learnt in computer vision and machine learning. We learned about how sometimes basic approaches work better than complicated approaches. Despite trying to use a smart segmentation algorithm, the relatively basic SMD model turned out to extract the best skin masks. We also realized the time constraints and difficulties of creating a dataset from scratch.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

Looking back, it would have been nice to have had a dataset already to work off of. Some Gestures were harder to classify in our live demo such as ‘’Hello’’ vs ”Thank You” since they only differ by a edge Although our classification system works quite well as has been demonstrated through tables and images, there is still a lot of scope for possible future work . For this project we have used Tensorflow Object Detection API which is used to train, develop and deploy the prototype system.
SSD has two components: a backbone model and SSD head. Backbone model usually is a pre-trained image classification network as a feature extractor. This is typically a network like ResNet trained on ImageNet from which the final fully connected classification layer has been removed. We are thus left with a deep neural network that is able to extract semantic meaning from the input image while preserving the spatial structure of the image albeit at a lower resolution. For ResNet34, the backbone results in a 256 7x7 feature maps for an input image. We will explain what feature and feature map are later on. The *SSD head* is just one or more convolutional layers added to this backbone and the outputs are interpreted as the bounding boxes and classes of objects in the spatial location of the final layers activations.

V. System Design and Flow

Using this prototype initially we can recognize almost 5 – 6 gesture including hello, yes, no, thank you and I love you. The prototype is very is to use and simple to execute and recognize the hand gestures with the help of system web cam frame.



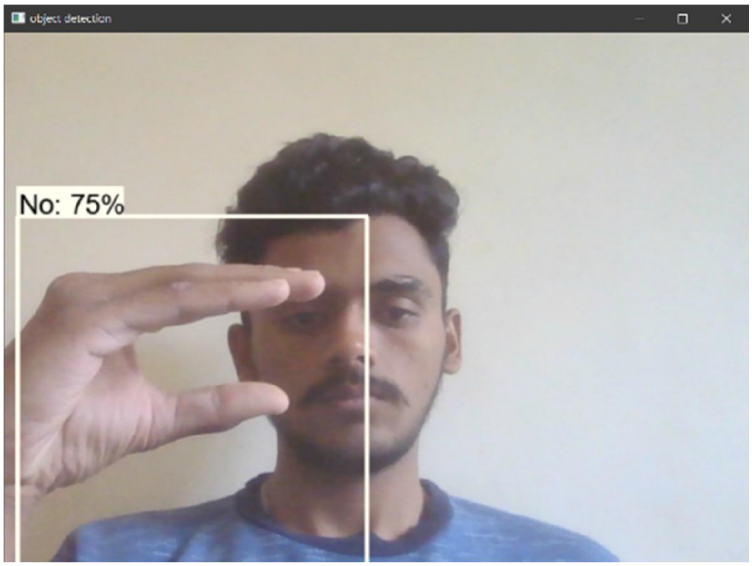
VI. Software Interface

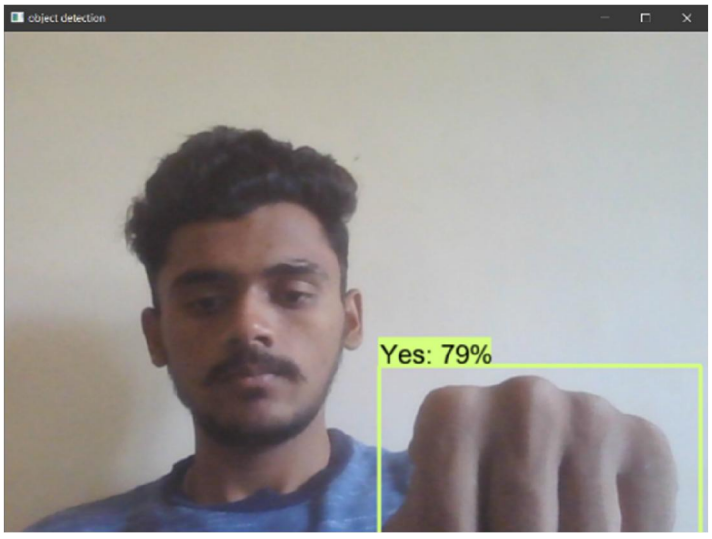
- **Jupyter Notebook :** Project Jupyter is a project and community whose goal is to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages. JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

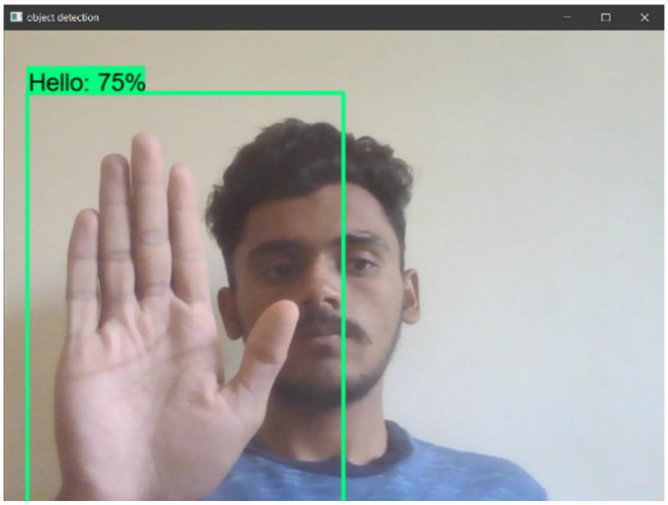
- **OpenCv :**OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

VII. PROJECT IMPLEMENTATION





IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252



VIII. Advantages & Disadvantages

- Easy to understand who don't know sign language.
- The model works well only in good lighting conditions.
- Plain background is needed for the model to detect with accuracy.

IX. Conclusion

In this Prototype, a functional real time vision based American Sign Language recognition for D&M people have been developed for Gestures and alphabets. Prediction has been improved after implementing SMD model in which we verify and predict symbols which are more similar to each other.

Acknowledgements

The completion of our project brings with it a sense of satisfaction, but it is never complete without those people who made it possible and whose constant support has crowned our efforts with success. One cannot even imagine our completion of the project without guidance and neither can we succeed without acknowledging it. It is a great pleasure that we acknowledge the enormous assistance and excellent co-operation to us by the respected personalities.

References

[1]. Akash. ASL Alphabet. url: https://www.kaggle.com/grassknoted/asl-alphabet. (accessed: 24.10.2018).
[2]. https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html
[3]. The Cognitive, Psychological and Cultural Impact of Communication Barrier on Deaf Adults". In: Journal of Communication Disorders, Deaf Studies Hearing Aids 4 (2 2016). doi: 10.4172/2375-4427.1000164.
[4]. https://en.wikipedia.org/wiki/Convolutional_neural_network
[5]. https://en.wikipedia.org/wiki/TensorFlow
[6]. Farnaz D. Notash and ElaheElhamki. "Comparing loneliness, depression and stress in students with hearing impaired and normal students studying in secondary schools of Tabriz". In: International Journal of Humanities and Cultural Studies February 2016 Special Issue (2016). issn: 2356-5926.