IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12620

SignOutLoud: Sign language Recognition

Dr. Shilpa Khedakar[1], Srushti Gaikwad[2], Anandi Gawade[3], Sakshi Sasane[4], Sakshi Memane[5]

Dept of Computer Engineering, Modern Education Society's College of Engineering, Pune[1-5]

Abstract: In everyday life, deaf people face many problems during simple task. One of the most difficult tasks for them is to communicate with other people. However only 0.25% of the total population use and practice Indian Sign Language. The proposed sign language Recognitions system aims to bridge the gap between non hearing and hearing population. This system is developed using Tensorflow, Tensorflow Object Detection API, Single Shot Detector (SSD). SSD is very popular algorithms for object detections amongst the other algorithms. These algorithms also provide high accuracy. This system converts the sign into text. India being more diverse and having many regional dialects still hasn't its own sign language. Overall accuracy of the model is 85%.

Keywords: Sign to text, Single Shot Detector, TensorFlow (TF), Tensorflow Object Detection API

I. INTRODUCTION

Communication plays an important role while communicating with other people where it is for a cause or exchanging your thoughts. Communication is very important for every human being and it also helps in the growth of the nations. When any one exchange their thoughts and opinion with other people, he can do better for themselves and also for the nation. Language is a form of communication that signal through many functions like posing question instead of deriving questions. Knowledge can be exchange with any person, whether they are deaf or mute. The problem occurs in the removable media. There are many different sign languages present today like ISL (Indian Sign Language), ASL (American Sign Language) but there is no fixed universal sign language is available.

Sign language play a very important role in exchanging information's from a person to a sewable deaf and hard of hearing people. Our India population is nearly 63 million with approximately 6.3% of the population is deaf. And according to WHO 25% of the world population will have hearing disability by 2050. As a result, it is very important to find a mechanism the enable the communication very easily. As a result, it is critical to know the mechanism that enables the disabled to communicate easily. They will enable the non-hearing population to present their thought to all the people and communicate with them easily. This would be very helpful to change the way they live the life and make their life simple and easier. Hearing loss has become a common problem worldwide.

World health Organization, 3.5 billion people (i.e one in five people) will have hearing loss by 2060 and they need hearing rehabilitation. So, there will be increase on the dependency on the sign language. Sign language will become the primary language for communication with different level of hearing impairment. Sign Language is a complete Language with their own syntax and grammar however its Properties Differ from the natural language. Every sign language has its own vocabulary which is lesser than the natural language. Therefore, the signer can use one sign to represent many different words. Similar to natural language sign language is also very diverse every region has its own dialects. Sign language is perform using hands and movement is the basic component of the sign language. Base on the movement it can be categorised into two parts: Static and Dynamic. Letters and number and one word can be represent using static hand gesture. Sentence are represented using dynamic hand gesture. Sign depends on the how we structure are hand to form a specific sign. Images can be captured of this gesture made by hand and can be use as a dataset for training. Static sign is more compared to the dynamic sign and those static signs are captured by images. For dynamic signs video stream is required to capture the sign.

Motivation –

- Lack of standard datasets
- The resolution of a sign is context-dependent.

Objective –

- Identify the symbolic expression through images.
- Creating our own words dataset.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12620

II.PROBLEM STATEMENT

Communication plays an important role while communicating with other people where it is for a cause or exchanging your thoughts. Sign language play a very important role in exchanging information's from a person to a sewable deaf and hard of hearing people. But the people having hearing and speech disability cannot communicate properly with hearing population as this people may not know the sign language. So, we proposed an interface by which sweable deaf and hard of hearing society can communicate via our system that would convert ISL (Indian Sign Language) into text.

III.LITERATURE SURVEY

1. Real Time Detection and Conversion of Gestures to Text and Speech to Sign System (IEEE)

Summary –
The proposed "Sign Language Translator" system aims to bridge the gap between deaf people and common people, and it is hoped that it will help Deaf-Mute people converse with people who are unfamiliar with ASL.
This paper proposes a model that can change over gesture-based communication into text and text/speech to sign language. As of now, this paper centres around the recognition of static signs of the American Sign Language (ASL) from pictures or frames from video.

Technology –
1. The goal of the system is to convert speech to American Sign Language by using an artificial neural network known as RNN (Recurrent Neural Network) with LSTM (Recurrent Neural Network - Long Short-Term Memory) trained by a Connectionist Temporal Classification (CTC) neural network
2. For gesture recognition this paper uses the SSD Mobilenet V2 model.

2. CNN Based Speech and Text Translation Using Sign Language (ICACCCN)

Summary –
Applications existence for sign language is necessary for dumb people so that they become able to talk easily with them who does not know sign language motions. Our aim to overcome this difference by using machine learning and neural networks to analyse sign language and translate it to text and voice.

Technology –

1. Image Processing with CNN
2. OpenCV

3. Hear Sign Language: A Real-Time End-to-End Sign Language Recognition System (IEEE)

Summary –
Sign language recognition (SLR) bridges the communication gap between the hearing-impaired and the ordinary people. However, existing SLR systems either cannot provide continuous recognition or suffer from low recognition accuracy due to the difficulty of sign segmentation and the insufficiency of capturing both finger and arm motions.

Technology –
1. Sign Speaker
2. Deep SLR to translate sign language into voices to help people "hear" sign language. Specifically.
3. Two armbands embedded with an IMU sensor and multi-channels EMG sensors are attached on the forearms to capture both coarse-grained arm movements and fine-grained finger motions
4. CNN

Accuracy – The average word error rate of continuous sentence recognition is 10.8 percent, and it takes less than 1.1s for detecting signals and recognizing a sentence with 4 sign words, validating the recognition efficiency and real-time ability of Deep SLR in real-world scenarios.

4. An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion (IEEE)

Summary –
In this paper, we propose a trainable deep learning network for isolated sign language recognition, which can effectively capture the spatiotemporal information using a small number of signs' frames.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12620

Additionally, the proposed approach outperformed the state-of-the-art techniques on the Argentinian sign language dataset LSA64.

Technology -
We propose a hierarchical sign learning module that comprises three networks: DMN, AMN, and SRN. Additionally, we propose a technique to extract key postures for handling the variations in the sign samples performed by different signers.
Accuracy – 15%

5. Sign language Recognition Using Machine Learning Algorithm (IRJET)

Summary –
This paper aims at extending a step forward in this field by collecting a dataset and then use various feature extraction techniques to extract useful information which is then input into various supervised learning techniques. Currently, they have reported fourfold cross validated results for the different approaches, and the difference from the previous work done can be attributed to the fact that in our fourfold I cross validation, the validation set Correspond to images of a person different from the persons in the training set.
Technology/Requirements –
1. Cross Validation,
2. Artificial Neural Network,
3. HU's moments,
4. Skin Segmentation,
5. SVM.

6. Sign Language to Voice Translator Using Tensorflow and TTS Algorithm (IEEE 2021)

Summary –
In this system, it helps to convert the sign language to voice with hand gestures understanding and capture the motion of hands. It is based on a Raspberry Pi with a camera module and is programmed in Python with the Open-Source Computer Vision (Open CV) library as a backend. Finally, the sign is identified by using Tensor flow algorithm and the outcome result as a voice using TTS algorithm. Open CV Python is implemented in this system. Various libraries are used in this system.

Technology/Requirements –
1. A Raspberry Pi and a module with a monitor the feature extraction algorithm was written in C++.
2. DCNN (Deep Convolutional Neural Network)
3. OpenCV
4. Tensorflow
5. CNN

Accuracy – Sign Recognition 90%

7. A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition (IEEE 2022)

Summary –
The aim of the present study is to provide insights on sign language recognition, focusing on mapping non-segmented video streams to glosses. For this task, two new sequence training criteria, known from the fields of speech and scene text recognition, are introduced

Technology/Requirements –
1. DNN
2. RGB+D dataset
3. Greek SL
4. CSLR training
5. CNN

Disadvantages of the existing system –
• The blur background recognition rate was less and the light conditions were mixed.
• Normal hands in hidden Markov language were not recognized.
• Hand occlusions were unaddressed.
• Recognize only close-range images, The greater the distance, the lower the brightness of the image.
• Template matching mechanism is very time consuming, so we need a better classification method existence exist.

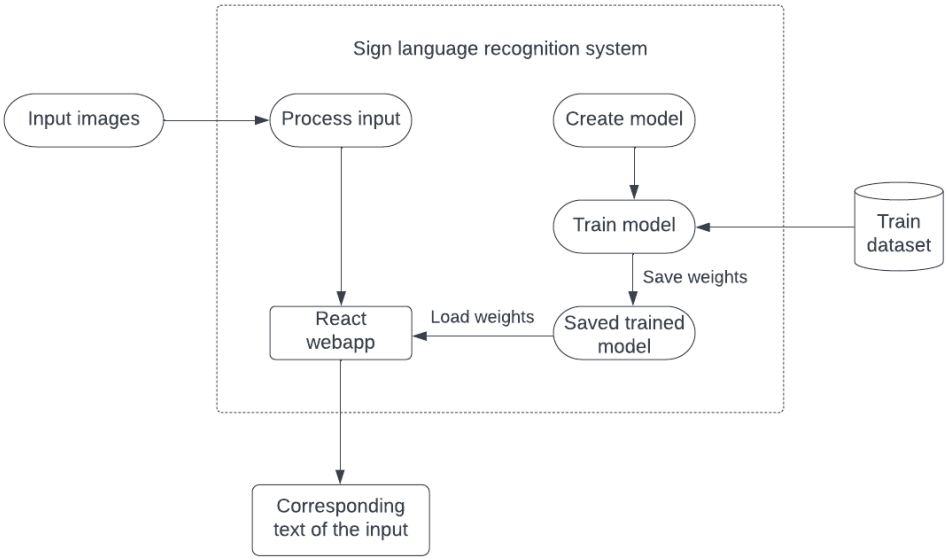
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12620

IV.PROPOSED METHODOLOGY



This online application is intended to solve the problems and difficulties faced by the people, including the difficulty in communication with deaf people. The users of this application can communicate with others using this web app and share their thoughts, feelings and experiences. This application bridges the gap between the users by allowing them to communicate easily with others.

Process:

- Gather the images from various users using OpenCV.
- Once you get correct and enough images, save them in the folder.
- Label those images using labelling tool which uses webcam feed and stores them in a folder.
- Add different variations like changing the background, changing the angel of camera, changing the hands, etc. to captured gestures.
- Run the code to split all captured gestures into training, validation and test data set.
- Then train that data using our model.
- Then save and export the model and it's related data in tf.js.
- Development of react application.
- Deploy and host the application on IBM Cloud.

The SignOutLoud system has a total of 2 modules

Frontend

React JS:

React is a JavaScript-based development library. React is library which is used to developed dynamic web application React is the most commonly used library for frontend development and it was first appeared in May 2013.React is not a language but a library for web development. React is one of the most popular libraries amongst all the library present today. The reason is easy creation of dynamic application, high performance, the component reusability, small learning curve and many more.

ReactJS, a JavaScript library for creating user interfaces, was utilised for the frontend development of Sign Out Loud. It is simpler to manage complicated UI elements with ReactJS' declarative approach to component creation. The header, footer, and main content are just a few of the parts that make up the frontend. Each component is designed to be reused on many website pages and has its own unique set of features. Sign Out Loud user interface is simple to use and straightforward. The user interface components were created and developed using a React UI framework. Users can use webcam to give the input to the system and it detect what user is exactly trying to expressed Based on the user's input, the frontend's output is presented in the form of text. The model offers the output by detections the user's gesture and responding appropriately using our model. In addition, we can add new and different signs to make future communication better and easier.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12620

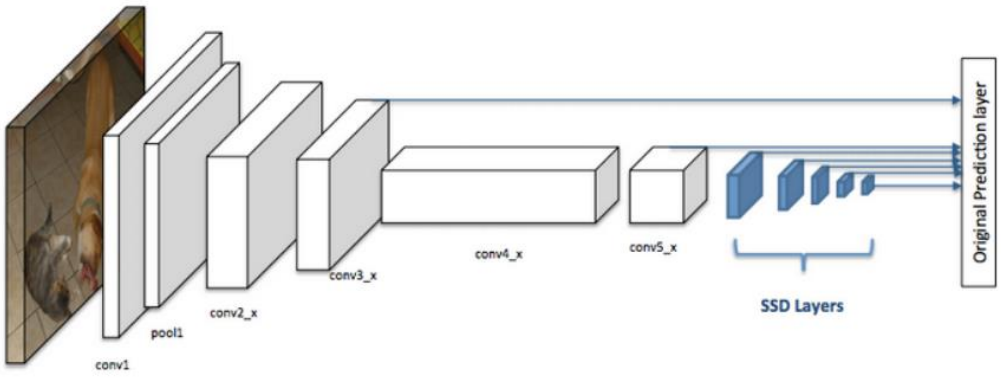
Backend

TensorFlow:

Tensorflow is one of the most popular library foe object detections. Tensorflow was developed by Google for deep learning application, machine learning application. Earlier TensorFlow was developed for Large numerical operation without deep learning in mind. In our project we have use Tensorflow object detection library for detection of the sign language. With the help of Tensorflow Library Zoo. We have made use of SSD_Mobilenet Model who has 22.2 COCO mAP and speed of 22.

SSD (Single Shot Detector):

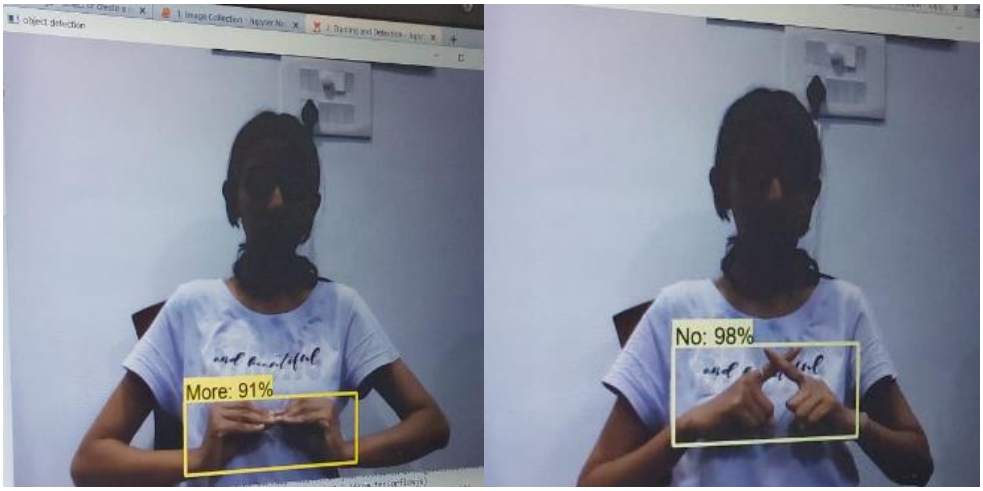
An SSD has two parts: the backbone model and the SSD head. A backbone model is usually a pre-trained image classification network as a feature extractor. This is typically a network such as ResNet trained on ImageNet from which the final fully connected classification layer has been removed. This leaves us with a deep neural network that is able to extract semantic meaning from the input image while preserving the spatial structure of the image, even at a lower resolution. For ResNet34, the backbone results in 256 7x7 feature maps for the input image. We will explain what functions and function map are later. The SSD header is just one or more convolutional layers added to this spine, and the outputs are interpreted as bounding boxes and feature classes in the spatial location of the final layer activations. In the image below, the first few layers (white boxes) are the backbone, the last few layers (blue boxes) represent the SSD head.



V. RESULT

We have collected the images for 16 sign which are Eat, Father, Mother, Pray, Thank You, I Love You, Help, Home, Time, Why, I, Sorry, Okay, Yes, More, No. Of each sign we have collected around 80 images. We have label all the images using labelling tools. And all the images were Spilt into train and test folder. With the help of train folder training was perform using SSD model.

Here are some screenshots of the result:

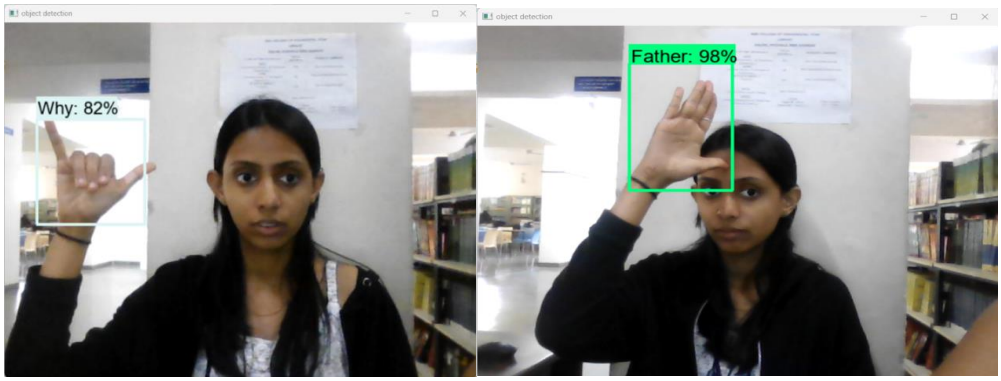


IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12620



The overall accuracy of model is 85%.

VI. CONCLUSION

In this work, we have implemented a system that identifies the symbolic expression through images. Currently, the system focuses on recognizing static Indian Sign Language (ISL) signs from images. Future research is expected to continuously recognize sign language with a dynamic hand gesture recognition system. For our system, we have created our own dataset which consists of words used commonly in day-to-day life. Our system was implemented using TensorFlow, Python, and the ReactJS library. We have used SSD model in our system which helps us in feature extraction and training our system. This system can recognize the words with an accuracy of 85% to 95% for each word. In the future, we decided to expand our dataset for more accuracy.

REFERENCES

[1] Saed Mian Qaisar, Sarah Niyazi, Abdulhamit Subasi, “Efficient isolated speech-to-character conversion based on Adaptive Rate Processing”; Procedia computing, St. 163, PP. 35–40, 2019.

[2] Prangon Das, Tanvir Ahmed , Md. Firoj Ali, “Static Hand Gesture Recognition for American Sign Language using Deep Convolutional Neural Network”, IEEE, Dhaka, Bangladesh, 2020

[3] L. Fernandes, P. Dalvi, A. Junnarkar and M. Bansode,” Bidirectional feature based on convolutional neural network Language Translation System”, Third International Conference on Smart systems and innovative technologies (ICSSIT), 2020.

[4] Vedant Puranik, Varun Gawande, Jash Gujarathi , Ayushi Patani, Tushar Rane, “Video-based Sign Language Recognition using Recurrent Neural Networks”; Asian Conference on Innovation in Technology (ASIANCON, Pune, India), 2022.

[5] Abhishek Wahane, Rishikesh Gadade, Aditya Hundekari, Aditya Khochare, Chudaman Sukte, “Real-Time Sign Language Recognition using Deep Learning Techniques”, IEEE 7th International conference for Convergence in Technology (I2CT),Pune, India, 2022

[6] Sunusi Bala Abdullahi, Kosin Chamnothai, ” American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach”, IEEE Access, Bangkok, Thailand, 2022

[7] Diksha Kumari, Dr R S Anand, “Static Alphabet American Sign Language Recognition using Convolutional Neural Networks”, international Conference on Advances in Computing, Communication and Materials (ICACCM), Uttarakhand, India, 2022

[8] Zhibo Wang, Tengda Zhao, Jinxin Ma, Hongkai Chen, Kaixin Liu, Huajie Shao, “Hear Sign Language: A Real-Time End-to-End Sign Language Recognition System”, IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 21, NO. 7, JULY 2022

[9] Hamzah Luqman, ” An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion”, IEEE, Saudi Arabia, 2022