iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.43797

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 VI June 2022- Available at www.ijraset.com

Real-Time Detection and Translation for Indian Sign Language using Motion and Speech Recognition

Swagata Katiyar[1], Rushali Mahajan[2], Nakshita Malhotra[3], Sanya Gupta[4]
Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology & Management Delhi, India

Abstract— Being able to communicate effectively is perhaps one of the most important life skills of all. Speaking is the primitive form of communication and it is what enables us to express ourselves. But life can get really difficult if one lacks the gift of auditory capability. These people communicate with something called sign language. Sign language is a distinctive yet exclusive language, which has been developed for deaf community to be a part of the common culture. In India, there is a large population who are dependent on this form of communication. However, due to the lack of awareness of sign language in our day-to-day lives, they feel isolated and disconnected with the world. Therefore, we have created a platform which can bridge the gap of this isolation and misunderstandings, using the concepts of Deep Learning. Sign-L is a sign language translator, which can translate actions to text and voice to actions through animation. Not just translations, but it also provides tutorials to learn sign language and increase the much-needed awareness among others as well.
Keywords— Sign language Translation, Communication, Social Services, Tutorial, Web Application

I. INTRODUCTION

Communication is a process of meaningful interaction among human beings, more specifically, it is the process by which meanings are perceived and reached among us. It helps us to understand people better, creates clarity of thoughts and expression, and also acts as a very important tool to educate people.

Linguistically, sign language is similar to any other language, facilitating deaf people to convey their thoughts or feelings through the movement of hands, making shapes, and using facial expressions. Not just deaf community, but people with disabilities like Autism, Apraxia of speech, Cerebral Palsy, and Down Syndrome also find sign language useful to communicate. Just like there exists thousands of verbal languages for hearing people, there also exists multiple sign languages for deaf community. There are somewhere between 138 and 300 different types of sign languages used around the globe today. But it is American Sign Language (ASL) which is the most common and globally accepted language. [11]

Despite having one of the largest populations of deaf people in the world, India still doesn't have a formally recognized system of sign language to call its own. Due to disintegration or lack of Indian Sign Language (ISL) resources, Indians find it inconvenient to learn or approach ISL. Our project Sign-L, solves this problem of unapproachability and insufficiency.

Implementing service integration is a major objective of successful application, which supports achieving its internal and external objectives at the same time. Through Sign-L, a well cumulative playlist, specifically designed for Indian Sign Language, will be made available to easily and effectively learn the language, as well as communication in real time through voice or text is also made possible through the same. Hence making Sign-L a single hub solution for all objectives revolving around the needs of ISL. Through growing advancement in Machine learning and its subset Deep Learning, deaf people will no longer feel like a minority from hearing cultures in everyday life. This will be a huge step closer to uniting and bringing humanity to its best.

II. BACKGROUND

In today's developing world, we ordinarily see the challenged, either as incapacitated or influential. Expressing uncertainty towards their capacities or being amazed by their achievements is the unbreakable existing norm. Given that India is a nation where ancient beliefs form an integral part of our culture and society, the unfounded assumptions concerning the impaired have opinionated the mindset of the community. From designing the educational system to enforcing laws for the society, the prevailing system takes account of the majority, putting the minorities to remain invisible to the available arrangements. [1]

Inclusion of people cannot take place till the time the programmes and schemes are confined to only a few checkboxes present on the forms or annexures. These notions regarding the impaired are deep-rooted in our lack of understanding of their lifestyle, beliefs and engagement with the surroundings. This inadequate information in connection with them often disregards their needs and desires and should be worked upon to get these people the fundamental necessities of living.

III.MOTIVATION

According to a census conducted by WHO, India has a community of nearly 63 million habitants who experience a Significant Auditory Impairment [8]. These individuals are often oppressed and get limited fundamental assistance including low-grade education, poor health facilities and minimum job opportunities. These individuals suffer the most because of the educational sector in India. Only 5% of these children get the basic schooling and 1% of the total of these communities achieve quality education. Furthermore, when these groups were questioned about their academic learnings, about 98% of them claimed that they have difficulties in comprehending as most of the teachers in special schools focus on oralism [9].

Presently with the introduction to COVID-19, it has become tougher for these individuals who would rely on oralism to read lips and recognise expressions to communicate while carrying masks and shields causing uncomfortable encounters. Despite the available assistive technology which has the potential to remove these barriers, individuals were compelled to navigate on online platforms such as Google Meet, Zoom and many others, without any additional support. [10]

The preliminary issues discussed above, highlights the importance of introducing new open-source platforms for reducing these exacerbating gaps. This project enlightens the above-mentioned issues and influences the scope of the project. The above data clarify the obstacles faced by the existing systems and narrows down the structure of the study.

IV.METHODOLOGY

A. System Architecture

The structure of the proposed work subdivides the system into two main portals, which have been created independently and can be accessed as per the user requirements.

1) *Translation Portal:* Consists of a speech recognition, or speech-to-sign animation gateway, which allows a machine or program to identify words spoken aloud and convert them into readable text. This is further rendered into sign language using animated avatar.
2) *Tutorial and Detection Portal:* This consists of a series of well gathered video tutorials through which users can learn the Indian Sign Language without having to navigate multiple platforms. Furthermore, it also consists of a motion recognition tab which is used for real time detection and pre-programmed recognition of gestures using sequences of images and videos achieved through AI.

B. Proposed Work

As mentioned above, this project is broadly divided in two sections which are responsible for direct or indirect translations. Motion Recognition and Speech Recognition are two main concepts used in Voice *Detection Panel* and *Voice Translation Panel* respectively.

The following are the steps required to obtain the desired goals.

1) *Motion Recognition:*

a) *Step 1:* Images are captured for multiple ISL conventions using OpenCV.
b) *Step 2:* In this step, operations are focused on different parameters of the given picture. Here, the captured image is cropped, filtered and adjusted according to the brightness and contrast. To acquire this efficiently, image enhancement, image cropping and image segmentation methods are used.
c) *Step 3***:** It consists of the sampling and labelling of the image. The following steps are considered to obtain high quality datasets for image training:

The image is converted from RGB format to a binary format.

After conversion, cropping of the image is done to obtain the key parts of the image.

Further, image enhancement takes place by focusing on a selected area to improve the quality and the information of original data.

Then the collected images are labelled using the python's LabelImg package, which is a graphical image annotation tool.

d) *Step 4:* Data is sent to get trained using ssd_mobnet which is a unified framework for object detection. It predicts the boundary boxes and their classes from feature maps in one single pass.
e) *Step 5***:** Finally, the model makes real time detections by recognizing multiple ISL gestures with high accuracy as shown in Fig 1.

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 VI June 2022- Available at www.ijraset.com

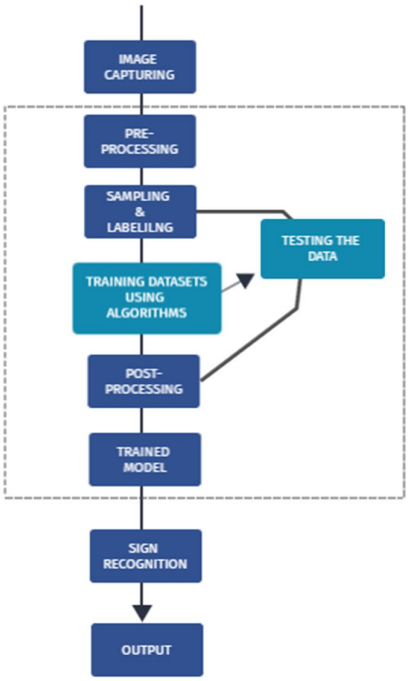


Fig. 1 Algorithm for Motion Recognition

2) *Speech recognition:*

a) *Step 1:* The end user inserts data in the form of voice, which is then converted into text by using the speech-recognition module NLTK. NLTK is one of the modules of the NLP algorithm, in which voice to text conversions are done with the use of a trained voice database.

b) *Step 2:* The speech recognizer first converts voice into text, which then split into words using word tokenizer. Then the translation module applies different rules that convert the tagged word/words into signs by means of grouping concepts.

c) *Step 3:* Lastly, the converted sign language is illustrated through a series of hand movements by an animated avatar (created through blender software) using our own database that have been incorporated within the system.

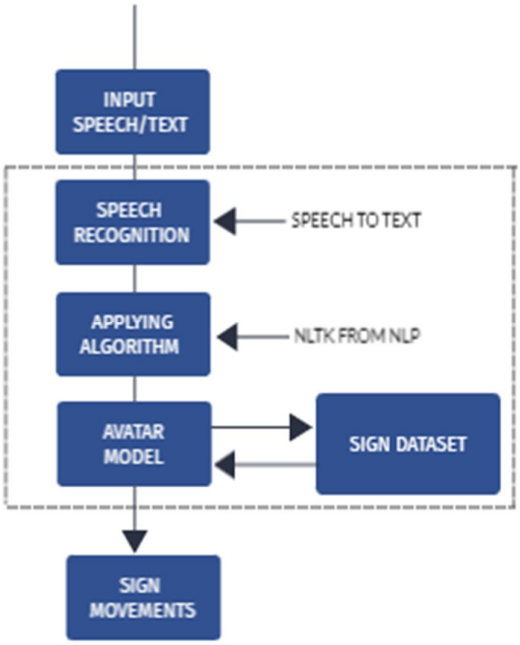


Fig. 2 Algorithm for Speech Recognition

V. ALGORITHMS

A. *CNN Algorithm*

A Convolutional Neural Network (CNN) is a Deep Learning algorithm which takes an input image and assigns importance to various aspects of it. Hence, helping in differentiating one object from another. In our proposed work, we applied a 2D CNN model with a tensor flow library for our *'Detection portal'* component. It consists of some convolution layers, each consisting of a pooling layer and an activation function. This architecture helps in training, extracting and generalizing the features of Indian Sign Language, to obtain high accuracy while making real time detections [12].

B. *NLP Algorithm*

Natural Language Processing (NLP) algorithm describes the interaction between human language and computers. The goal of using NLP in our project is to make the systems understand unstructured texts and retrieve meaningful pieces of information from it. NLTK, or Natural Language Toolkit, is a Python package that can be used for NLP. After installation of NLTK library and data packages, tokenization technique has been used. This technique helps us to break the given texts into smaller units which are also called tokens [13].

VI. RESULTS

Sign-L is a working model of Indian Sign language detection using Deep Learning and some dependencies of python like TensorFlow, Blender (Animation), OpenCV, CRF Tokenizer, Django (web framework) and many more. It is built within the standards of ISL, which is further divided into data acquisition and classification.

The dataset consists of a healthy mixture of images related to alphabets, numbers and many other important words. The training accuracy and training loss were obtained as *94.65%* and *0.0259* respectively. And the test accuracy was *94.62%*. The number of steps used to train the respective model were about *20,000*.

The collection of images with hands and its gestures were obtained from a web camera. The main advantage of using an inbuilt camera is that it removes the need for sensors in sensory gloves and reduces the cost of building the system. Also, since the web cameras are quite cheap and are available in almost all laptops, it was a convenient resource to use.

VII. CONCLUSION AND FUTURE SCOPE

Gestures are deeply rooted in the deaf community and their lifestyle. However, the majority are not familiar with this nonverbal form of communication which results in a linguistic rift. To avoid such circumstances, the idea of Sign-L was introduced.

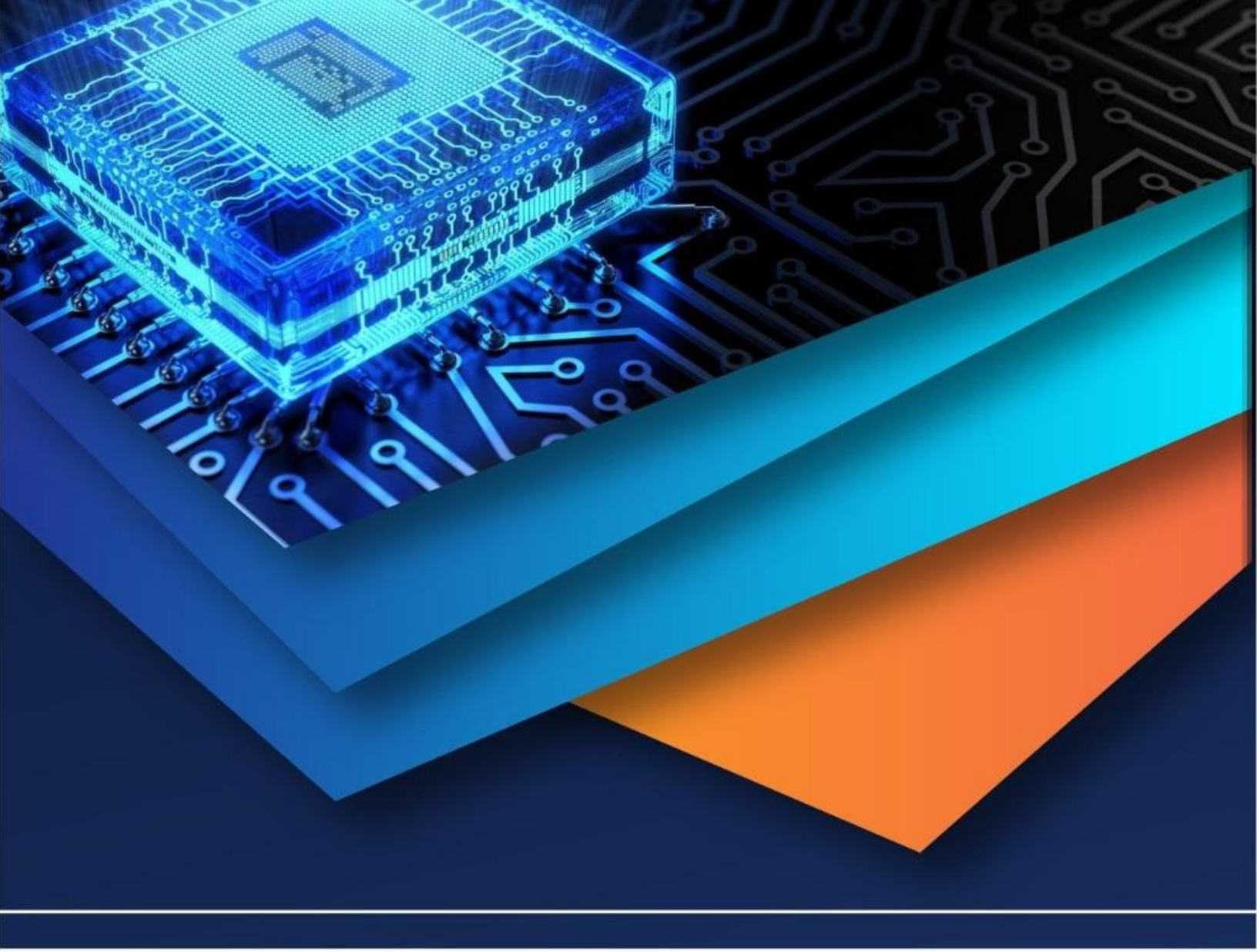
Mobile phones in the future are expected to be more closely embedded in our day-to-day lives than ever before. They still are more convenient to be used for most of the applications. It is obvious that platforms like Sign-L will be more productive and convenient when embedded in smartphones, making real time translations handier and on the go. Thus, making the transition into a mobile app would be our prioritized forthcoming work.

Due to the obscure resources of Indian Sign Language, we worked on ISL based datasets in a hope to promote this form of communication among Indians [5]. But since ASL is considered as a standard sign language, we are also planning to incorporate ASL in our platform. Hence our datasets will be increased significantly, making it more versatile, accurate and useful.

Hopefully, our application will serve to forge stronger, more trusted bonds between deaf-mute and the general community. To overcome possible issues, the system's components have been designed and are being tested individually and all together as well. To meet the individual demands for each component and to solve any potential obstacles, the application will be assessed and its accuracy will be examined and compared to those of other systems.

REFERENCES

[1] https://www.researchgate.net/publication/262187093_Sign_language_recognition_State_of_the_art
[2] https://www.irojournals.com/iroiip/V2/I2/01.pdf
[3] https://www.ijert.org/a-review-paper-on-sign-language-recognition-for-the-deaf-and-dumb
[4] https://www.upgrad.com/blog/top-dimensionality-reduction-techniques-for-machine-learning/
[5] https://www.academia.edu/75517462/Captioning_and_Indian_Sign_Language_as_Accessibility_Tools_in_Universal_Design
[6] https://www.academia.edu/53205170/Vision_Based_Hand_Gesture_Recognition_Using_Fourier_Descriptor_for_Indian_Sign_Language
[7] https://www.businessinsider.in/india/news/hidden-behind-masks-people-with-speech-hearing-disabilities-struggle-to-communicate-in-covid-19-times/articleshow/76441910.cms
[8] https://en.wikipedia.org/wiki/Deafness_in_India
[9] https://timesofindia.indiatimes.com/city/lucknow/govts-deaf-mute-approach-turns-them-into-handicapped/articleshow/17210135.cm
[10] https://www.indiatoday.in/education-today/featurephilia/story/deaf-children-hearing-impaired-dont-have-college-options-teach-organisation-solving-problem-html-1313347-2018-08-24
[11] https://en.wikipedia.org/wiki/List_of_sign_languages
[12] https://towardsdatascience.com/a-comprehensive-guide-to-conAvolutional-neural-networks-the-eli5-way-3bd2b1164a53
[13] https://en.wikipedia.org/wiki/Natural_language_processing



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)