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.2022.43997

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

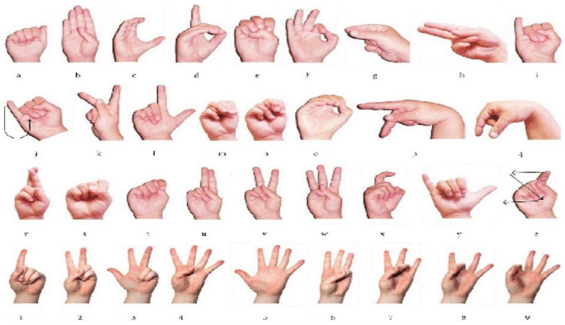
Sign Language Detection Using Hand Gestures

Naadem Divya, Kranthi Kumar. K, Sathvik Vemulam, Rohan Rao Regulapatim, Akshay Chiluveru
[1,2]*Associate Professor,* [3,4,5]*B. Tech IV year , IT Department*
Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad

Abstract- Some of the major problems faced by a person who are unable to speak are, they cannot express their emotion and they are unable to use (Artificial Intelligence) like google assistance, or Apple's SIRI etc because all those apps are based on voice controlling. Hand gestures are one of the nonverbal communication strategies used in sign language. People who are deaf or hard of hearing are the ones who use it the most to communicate with one another and with others. Various sign language systems have been developed by several firms across the world, however they are neither customizable nor cost-effective for end users. We developed an app that recognises pre-defined American signed language using hand gestures (ASL).Our application will have two main featuresThe motion has been recognised, and the corresponding letter has been shown. The second feature is to display a meaningful word. 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. Sign language contains special rules and grammar's for expressing effectively. In a range of applications, such as human-computer interfaces, multimedia, and security, gesture recognition is becoming more significant. Our programme would be a user-friendly, comprehensive system.

I. INTRODUCTION

One can share the thoughts with the other person through communication but it's very difficult for the people who cannot talk & who cannot hear. Here comes the solution, Sign language is the best alternative bridge the gap of communication for dumb and deaf. Each gesture has certain meaning in Sign language. Even with the very basic elements, complex meanings can also be expressed and explained. It developed as a standard gesture-based language for deaf and dumb people to communicate.This type of non-verbal language helps dumb and deaf to communicate among themselves and even with the common people. For expressing effectively, Sign language contains certain protocol to abide by. Gesture recognition is becoming prominent in numerous variety of applications like human interface communication, multimedia and security.



Various systems for sign language are developed by different manufacturers but they are neither efficient nor serving the needs of common man. It is out of reach for the average person's budget.Our method of approach involves implementing such application which follows pre-defined American sign language(ASL) through hand gestures. Our application will satisfy the needs of dumb and deaf, technically sophisticated with comprehensive features. We developed that sort of application which detects hand gestures with the help of Signs. Hands come in a variety of forms and orientations, depending on who you ask. Depending on their personal style, everyone employs a variety of gesture postures. We need an application which can handle that complexity in the input.

Hand motion non-linearity is a severe issue that must be addressed. One simple solution to solve this would be collecting metadata and using on top of it. It is nothing about information about an information. It tells what type of data is stored rather than revealing the data itself. It recognises gestures by analysing metadata content information in photographs. This procedure combines two distinct tasks: feature extraction and classification. Before any gesture can be recognised, picture characteristics must be retrieve. One of the categorization strategies should be used after the extraction of such attributes. The main issue here will be determining how to extract the fundamental characteristics and imposing them for categorization. A large number of qualities are required to identify and recognise someone.

CNN is a deep learning neural network that stands for convolutional neural network. It has the ability to extract characteristics. For classification, fully connected layers might be used. For a system's memory and computational complexity to be reduced, CNN combines these two procedures. It gives you the best results. It can comprehend the pictures' complicated non-linear connections. Hand Gestures for Detecting Sign Language 2 As a result, the ideal strategy to handle this problem is to use a CNN-based approach. This project emphasizes on single gesture recognition, own gesture creation, creating meaningful sentences and can export those sentences into text file.

II. PROBLEM STATEMENT

Given a hand gesture, implementing such an application which detects pre-defined American sign language (ASL) in a real time through hand gestures and providing facility for the user to be able to store the result of the character detected in a txt file, also allowing such users to build their customized gesture so that the problems faced by persons who aren't able to talk vocally can be accommodated with technological assistance and the barrier of expressing can be overshadowed.

III. PROPOSED SYSTEM

Our programme recognises established American sign language using hand gestures. PyQt5 module is used to build this system which makes it user friendly. A CNN-based method was employed in the construction of our application. It is capable of comprehending the visuals' intricate and non-linear interactions. Application comprises of user guide like a video at start which explains the step by step procedure for using the application. It will tell how to create a gesture and the procedure that needs to be followed. It also helps in knowing about the detailed information of Scanning single gesture, framing sentences and exporting to text file.

A. Create Gesture

With the help of text box, user can give his own hand gesture to the system. At the bottom of the screen, the user must type in the motion's name. For further processing, since this gesture is customized, it needs to be stored inside a folder.

B. Scan Single Gesture

A gesture scanner is available at front of the end user in which user has to show gestures with his hands. Based on the output of pre-processed module, a user can see associated label assigned for it on the output window screen.

C. Scan Sentence

A user can select a delimiter. Until the delimiter is encountered, every gesture which is scanned will be appended with the previous results which will form a stream of meaningful words and sentences.

D. Exporting file

The findings of the scanned characters can be exported to a separate text file by the user.

IV. ARCHITECTURE

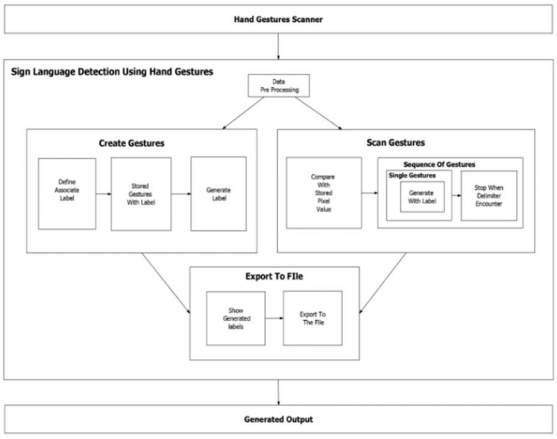


Fig 2: Architecture

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

Conv-Net / CNN is an in-depth learning approach that focuses on image acquisition such as input and usage of weights and dependencies that can be studied for different parts of it. You can tell the difference between the two. Other classification algorithms require more pre-processing than ConvNet. Filters are created by hand using simple methods, but with enough training, ConvNets can learn these filters,features.

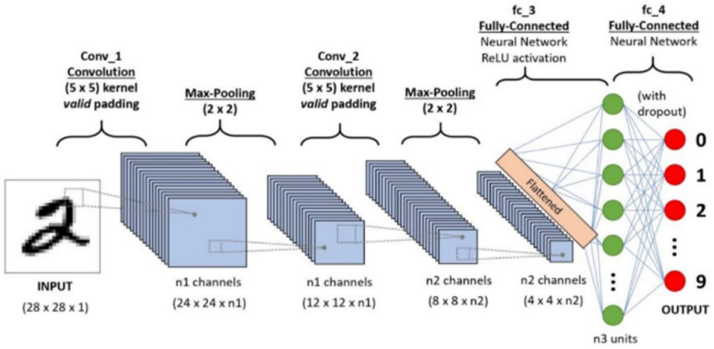


Fig 3: CNN Architecture

A ConvNet's architecture is analogous to the neuronal connection network in the human brain. It is based on the structure of the visual cortex. Each neuron responds only to stimuli located in the receptor field, which is a small part of the visual field. A number of similar fields can be stacked on top of each other to encompass the full visual field.

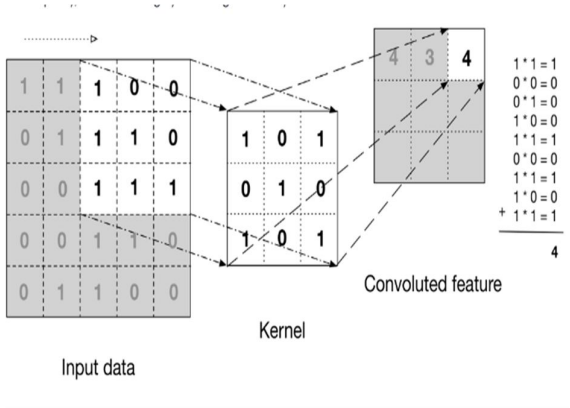


Fig 4: Convolution layer for grayscale images

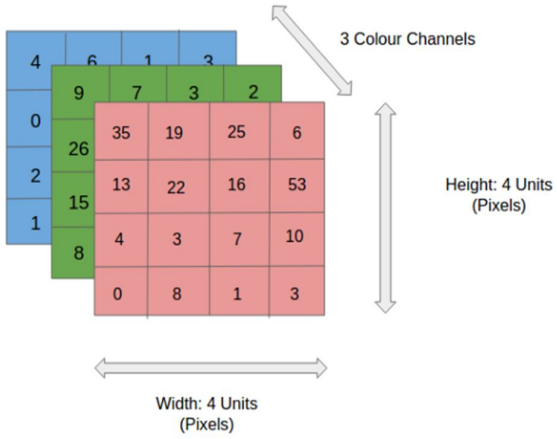


Fig 5: Convolution layer for color images

The aggregation level is responsible for reducing the size of the space of complex objects. This helps to reduce processing power. Pooling may be divided into two types: average pooling and maximum pooling. Average Pooling is the process of averaging all of the values in the Kernel's region of the image. Max Pooling calculates the maximum value of a pixel from a region of the image covered.

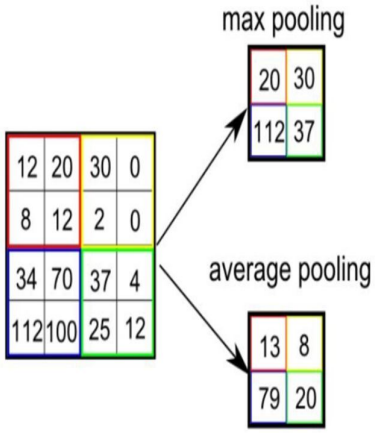
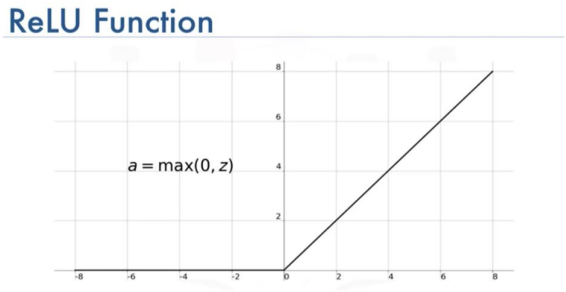


Fig 6: Pooling Methods

A. Activation Function

A modified line execution function, or ReLU, is part of a line function that directly outputs an input if it is negative, and outputs zero if the input is positive. Because learning is easier and more effective, it has become the default activation feature for many types of neural networks.

The modified linear activation function solves the problem of the disappearance gradient and allows the models to train faster and work better.



V. METHODOLOGY



Fig 7: Methodology

A. Input Data and Training Data

The camera is used to collect pictures for the model's training and validation. The persons in front of the webcam make the gestures. In the input images, one hand is expected to be present; motions are done with the right hand, with the palm towards the camera and the matching hand nearly vertical. If the backdrop is less complex and the contrast is strong on the right hand, the recognition process will be less complicated and more efficient. We presume the backdrop of the image is less complicated and more homogeneous.

B. Pre-processing

To reduce computer difficulties, a little amount of pre-processing was applied to the dataset. It also aids in the improvement of efficiency. Using the background subtraction approach, the photos' backdrop was eliminated. Z. ZivKovic has proposed it. It is mostly based on the K-gaussian distribution, which finds the optimal gaussian distribution for each pixel, allowing for improved adaptation under a range of lighting conditions. After removing the whole backdrop, only the picture of a hand remains. Grayscale conversion was performed on these photos. There is only one colour channel in them. CNN will be able to learn and adjust more quickly as a result of this. After that, morphological erosion was employed. The noise was then reduced using a median filter. It is the only signal processing component that decreases noise.

C. CNN Configuration

In this study, CNN was used to determine hand movements. The design consists of two layers of convolution, two layers of maximum integration, two fully interconnected layers, and one source layer. To minimize readjustment, the network includes three options. The first coil layer consists of 64 separate filters with a core size of 3x3. This layer's activation function is Rectified Linear Unit. On top of that, it's utilised to introduce non-linearity, and it's been shown to outperform other activation functions like tanh and sigmoid. Because it is an input layer, we must provide the input size. The stride has been restored to its previous value. The input shape is 50x50x1, meaning that this network should receive a grayscale image of 50x50 pixels. This layer produces map features before passing them on to the next layer. A max pooling layer with a pool size of 2x2 is then added to the CNN, which extracts the maximum value from a 2x2 window. Because the pooling layer only takes the greatest value and discards the remainder, the representation of spatial size is gradually decreased. Because it only chooses the most relevant information, this layer aids the network in better understanding and deciphering the pictures.

D. System Implementation

Python programming language is used to implement and develop this amicable deaf and dumb system. python IDE Spyder is used for developing and running the code. Availed the facility of Keras library in order to build the CNN classifier. For visualization, Matplotlib is used. It contains information such as the model's accuracy, loss values, and confusion matrix. NumPy was utilised to perform array operations. On top of the dataset, there are two steps to the training process.

1) Training using a Basis Dataset: The model was trained with a base dataset acquired after pre-processing.

2) Expanded Dataset Training: In this stage, the dataset was expanded. Data augmentation is a means of expanding the number of data chunks by employing techniques such as zooming, shearing, rotating, and flipping. This strategy not only adds to the data, but also to the dataset's diversity, which is required for CNN to learn localised picture differences.

VI. OUTPUT SCREENS

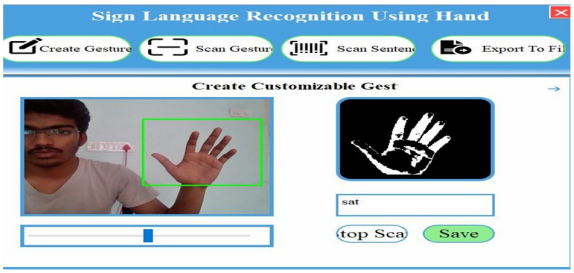


Fig 8: Creating gesture

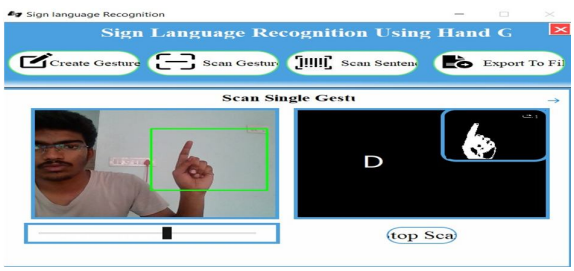


Fig 9: Scan Gesture

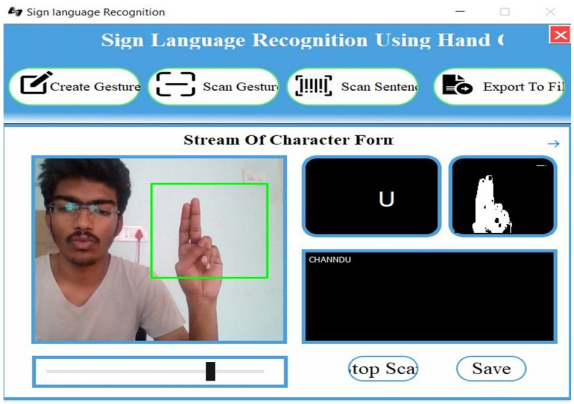


Fig 10: Formation of sentence



Fig 11: Exporting to a file

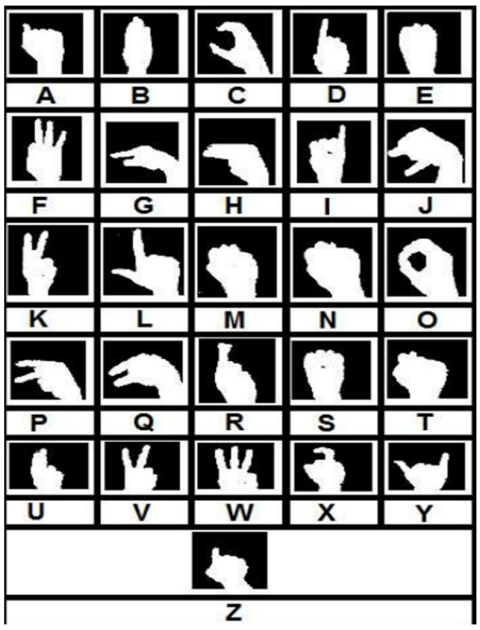


Fig 12: All gestures in a window

VII. CONCLUSION

Our application will be the front runner and act as solution for some of the major problems faced by the differently abled people in terms of communication.

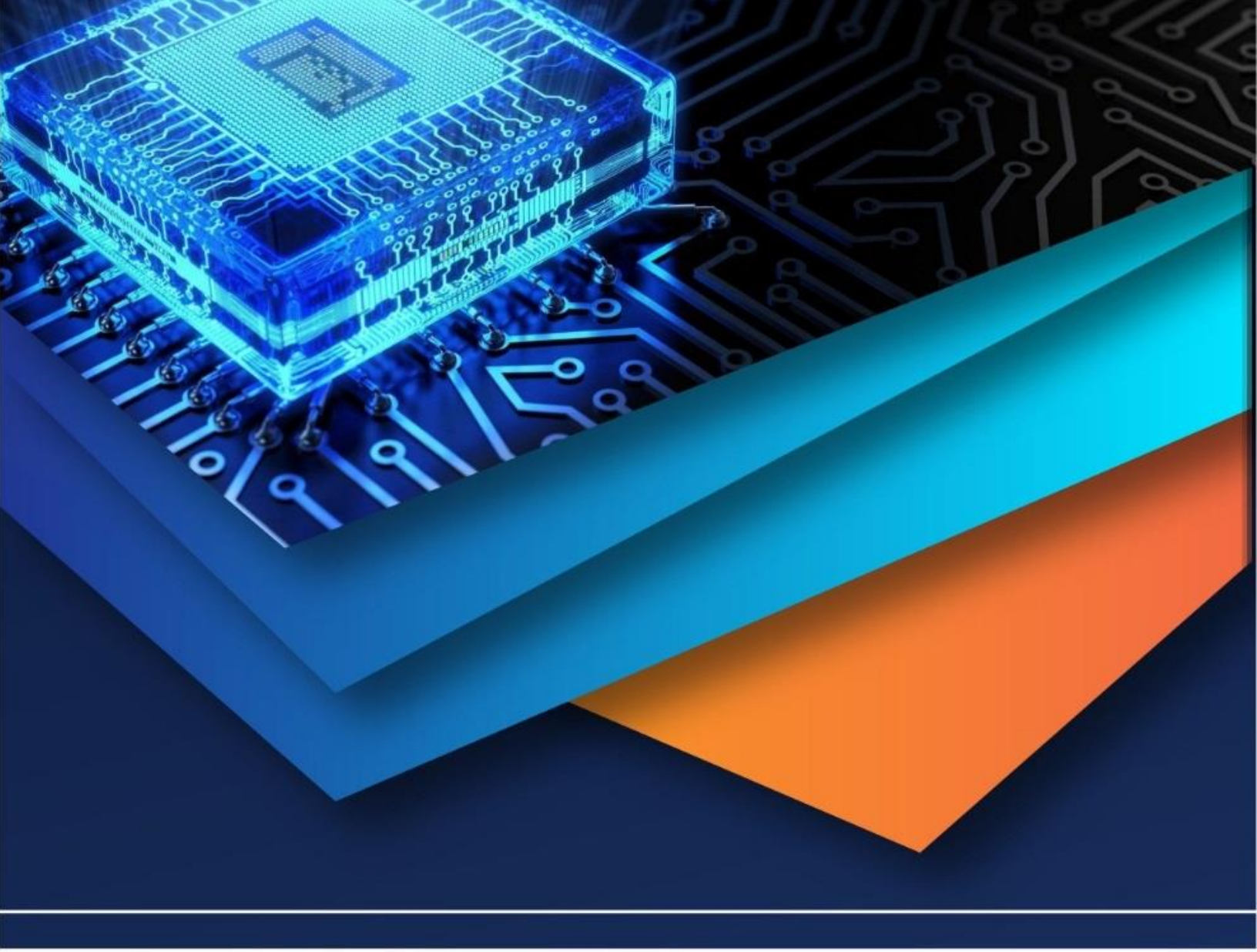
This application helps in serving the person who wants to learn and talk in sign languages. Availing the features of this application, a person can adapt quickly for various gestures and their meaning according to ASL standards. They can imbibe and inject in a faster pace knowing what alphabet is assigned to which gesture. Custom gesture facility is coupled with different features like it can provide a facility for sentence formation. There are many illiterates who can use this, user need not to be literate, if they know how to show the gesture actions, In a short period of time, the given character will appear on the screenx.

Furthermore, an export to file module is included with TTS (Text-To-Speech) support, which means that whatever phrase the user forms will be able to listen to it and then swiftly export while also seeing what gestures the applicant used throughout the production of the sentence.

REFERENCES

[1] NIDCD, “american sign language”, 2017 https://www.nidcd.nih.gov/health/americansign-language,
[2] Shobhit Agarwal, “What are some problems faced by deaf and dumb people while using todays common tech like phones and PCs”, 2017 [Online]. Available: https://www.quora.com/What-are-some-problems-faced-by-deaf-and-dumb-peoplewhileusing-todays-common-tech-like-phones-and-PCs, [Accessed April 06, 2019]
[3] M. Ibrahim, “Sign Language Translation via Image Processing”, https://www.kics.edu.pk/project/startup/203
[4] NAD, “American sign language-community and culture frequently asked questions”, 2017 https://www.nad.org/resources/american-sign-language/community-andculturefrequently-asked-question
[5] Sanil Jain and K.V.Sameer Raja, “Indian Sign Language Character Recognition”, https://cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com



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)