IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

Sign Language Recognition Using Hand Gestures

Khateeja Ambareen[1], Rakshitha S[2], Sahana KJ[3], Varshini V[4], Yashashwini V[5]

Assistant Professor, Department of CSE, GSSSIETW, City, Mysuru, India[1]

UG Student, Department of CSE, GSSSIETW, Mysuru, India[2-5]

Abstract: Deep learning models, namely, LSTM and GRU (feedback-based learning models), are used to recognize signs from isolated Indian Sign Language (ISL) video frames. The four different sequential combinations of LSTM and GRU (as there are two layers of LSTM and two layers of GRU) were used with our own dataset, IISL2020. The predominant means of communication is speech; however, there are persons whose speaking or hearing abilities are impaired. Communication presents a significant barrier for persons with such disabilities. The use of deep learning methods can help to reduce communication barriers. This paper proposes a deep learning-based model that detects and recognizes the words from a person's gestures. Deep learning models, namely, LSTM and GRU (feedback-based learning models), are used to recognize signs from isolated Indian Sign Language (ISL) video frames. The four different sequential combinations of LSTM and GRU (as there are two layers of LSTM and two layers of GRU) were used with our own dataset, IISL2020. The proposed model, consisting of a single layer of LSTM followed by GRU, achieves around 97% accuracy over 11 different signs. This method may help persons who are unaware of sign language to communicate with persons whose speech or hearing is impaired.

Keywords: LSTM and GRU (feedback-based learning models), Deep learning models, Indian Sign Language (ISL).

I. INTRODUCTION

Sign language is very essential for deaf-mute people, to communicate both with normal people and with themselves, is still getting less attention from the normal people. Hand gesture is one of the methods used in sign language for non-verbal communication. It is most commonly used by deaf & dumb people who have hearing or talking disorders to communicate among themselves or with normal people. HCI finds its applications in augmented reality systems, facial recognition systems and also hand-gesture recognition system. This project falls under the domain of HCI and aims towards recognizing various alphabets (a-z) and digits (0-9) of the ISL family. Hand gesture recognition is a challenging task, particularly the ISL recognition is complicated due to its usage of both the hands. In the past many works have been performed in this respect using sensors (like glove sensor) and other image processing techniques (like edge detection technique, Hough Transform etc.) but were unable to achieve satisfactory results.

II. REVIEW OF OTHER METHODS

1. Sign Language Recognition Using Template Matching Technique- Soma Shrenika, Myneni Madhu Bala

This system bridges the imbalances between deaf people and normal people without any requirement of an intermediate translator. It achieves the objective of conversion of gestures to text. The process in brief includes acquiring images using a camera. Then performing pre-processing steps on the image, that is, convert the acquired image, which is in RGB model to gray scale image. Later, track the edges by using canny edge detection algorithm. Finally, detecting the pattern using template-matching algorithm, this outputs the result as text. This paper is to help and serve the deaf of our society to communicate with normal people. Here the implementation of the system is using image-processing techniques. This system is for people who cannot use gloves, sensors and other highly refined equipment. First, acquire image with a camera. Then convert it to gray scale image for further processing. Edge detection algorithm was used to detect the sign in the image. There, the process includes removal of noise and other less important data and applying smoothing algorithm to image, finding gradient magnitude followed by tracking the edges by hysteresis. The last step is displaying the sign alphabet. In future, we can develop a system that is two-way system where, conversion of sign to text and text to sign is possible. Developing a system, where interpretation involves dynamic gestures. Implementation can extend to mobile phones.

2. Real-Time Recognition Of Indian Sign Language- Muthu Mariappan H, Dr Gomathi V

Sign languages are a visual representation of thoughts through hand gestures, facial expressions, and body movements. Sign Languages also have several variants, such as American Sign Language (ASL), Argentinean Sign Language (LSA),

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

British Sign Language (BSL) and ISL. The hearing and speech impaired people prefer the sign language, which is mostly used in their region. Moreover, in India, there is no universal sign language. Though there exist many sign languages, the normal people do not know about sign languages. Hence communicating with deaf and dumb people becomes more complex. The proposed system has a camera unit for capturing the gestures of the hearing and speech impaired people. The real-time sign language recognition system was designed as a portable unit for more convenience of the users. The raw videos taken in a dynamic background is given as an input to the system. The image frames are resized to maintain the equality among all the videos. OpenCV (Open Source Library for Computer Vision) is used for feature extraction and video classification. The system for recognizing real-time Indian Sign Language (ISL) portrays an impressive role in enhancing casual communication among people with hearing disabilities and normal persons. Though FCM is efficient, it requires more computation time than the others. Also, for high dimensionality datasets, most of the traditional algorithms suffer. Hence it is planned to extend the system by combining Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to capture the spatial and temporal features. In future work, more words will be added to the system.

3. Dynamic Hand Gesture Recognition for Indian Sign Language- Bhoomi Lodaya, Dr. Narendra Patel, Dr. Hemant Vasava

Sign Language is hard to understand for the normal person. According to WHO around 466 million people in world are dumb or deaf i.e., approximately 5% of total's world population. [1] Sign Language is a language like any other globally recognized language. It is a nonverbal means of communication used by the deaf and mute to communicate amongst each other and with others. It allows people to express themselves and understand each other without speaking. There are many different sign languages, for example Indian Sign Language, American Sign Language, Chinese Sign Language etc. In the current scenario American Sign Language is the most used language among all. We are proposing a simple system to convert the Indian Sign Language to text by identifying the sign using CNN. The technology is evolved and has potential to help the dumb and deaf people. We propose a model that can recognize sign language using the Convolutional network as in Fig 1. The user can recognize sign language using a smartphone camera and the predicted alphabets will be visible on the real time basis. In addition to that we can also make the model available for different sign languages. To develop this model we used tensorflow library and keras library. For this we created the dataset using the Video and extracted images or frames using OpenCV library. This system helps many dumb and deaf people. Though this CNN model is efficient, it needs more computations for building and loading model. The proposed model takes input image and recognizes the sign language. We proposed the model with an accuracy of 98.34%.

4. Deep sign: Sign Language Detection and Recognition Using Deep Learning -Deep Kothadiya, Chintan Bhatt, Krenil Sapariya, Kevin Patel, Ana-Belén Gil-González and Juan M. Corchado

Sign language can be divided into two main categories: static and dynamic. Static signs are steady hand and face gestures, while dynamic signs are further divided into isolated signs and continuous signs. The aim of the present study was to develop a system that recognizes sign language, and that can be used offline. A vision-based approach was developed to obtain the data from the signer. One of the main features of this study is the ability of the system to identify and recognize the words included in IISL2020 (our customized dataset). The IISL2020 dataset consists of 11 words; for each word, there are about 1100 video samples of 16 research participants, including males and females. The IISL2020 dataset was created considering natural conditions—without extra brightness, orientation, background adjustments, gloves, etc. Unlike most sign language datasets, the authors did not use any external help, such as sensors or smart gloves to detect hand movements. The GRU and LSTM methods are used because of their ability to remember previous inputs with the help of gates. The corresponding word is learned on the basis of the previous activation and current input feature, and the weights are adjusted accordingly. The training dataset is further split into 80:20 ratios, sorting the data into training and validation datasets. The validation dataset can be useful to test the learning achieved by the model after each epoch.

III. METHODOLOGY

Data Pre-Processing – In this module, based on the object detected in front of the camera its binary images is being populated. Meaning the object will be filled with solid white and background will be filled with solid black. Based on the pixel's regions, their numerical value in range of either 0 or 1 is being given to next process for modules. ➢Scan Single Gesture – A gesture scanner will be available in front of the end user where the user will have to do a hand gesture. Based on Pre-Processed module output, a user shall be able to see associated label assigned for each hand gestures, based on the predefined American Sign Language (ASL) standard inside the output window screen. Create gesture –A user will give a desired hand gesture as an input to the system with the text box available at the bottom of the screen where the user needs to type whatever he/she desires to associate that gesture with. This customize gesture will then be stored for future

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

purposes and will be detected in the upcoming time. Formation of a sentence – A user will be able to select a delimiter and until that delimiter is encountered every scanned gesture character will be appended with the previous results forming a stream of meaning-full words and sentences. Exporting – A user would be able to export the results of the scanned character into an ASCII standard textual file format. Testing is a process of executing a program to ensure that defined input will produce actual results that agree with required outputs. In developing a software project, error can be initiated at any stage during the development. For each phase of the software development cycle there are different techniques for detecting and elimination errors that originate in that phase. However, some errors will reflect in the code. Testing performs a very crucial role for quality assurance and for ensuring the reliabilities of the software. The quality of the system depends on its design, development, testing and implementation. Weaknesses in any of these areas will seriously affect the quality and therefore value of the system to its users. Once the code has been generated, testing of the modules begins implementation ends with formal tests. Functional Requirement defines a function of a software system or its component. A function is described as a set of inputs, the behaviour and outputs. Functional Requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what the system is supposed to accomplish. Behavioural requirements describing all the cases where the system uses the functional requirements are captured in use cases. Functional requirements are supported by non-functional requirements, which impose constraints on the design or implementation such as performance requirements, security or reliability. Non-Functional Requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. This should be contrasted with functional requirements that define specific behaviour or functions. In general, functional requirements define what a system is supposed to do whereas non-functional requirements define how a system is supposed to be. Non-functional requirements are often called qualities of a system. Other terms for non-functional requirements are “constraints”, “quality attributes”, “quality goals” and “quality of service requirements” and “non-behavioral requirements”. Security: the program should not harm user’s data, corrupt files or cause any viral activity in the event of program crash. The application should not allow users to access the codes of other users for modification. Ease of Use: A reasonably good interface has to be provided. The user should be comfortable in using the application. The application should not confuse the user with too many options in a single screen.

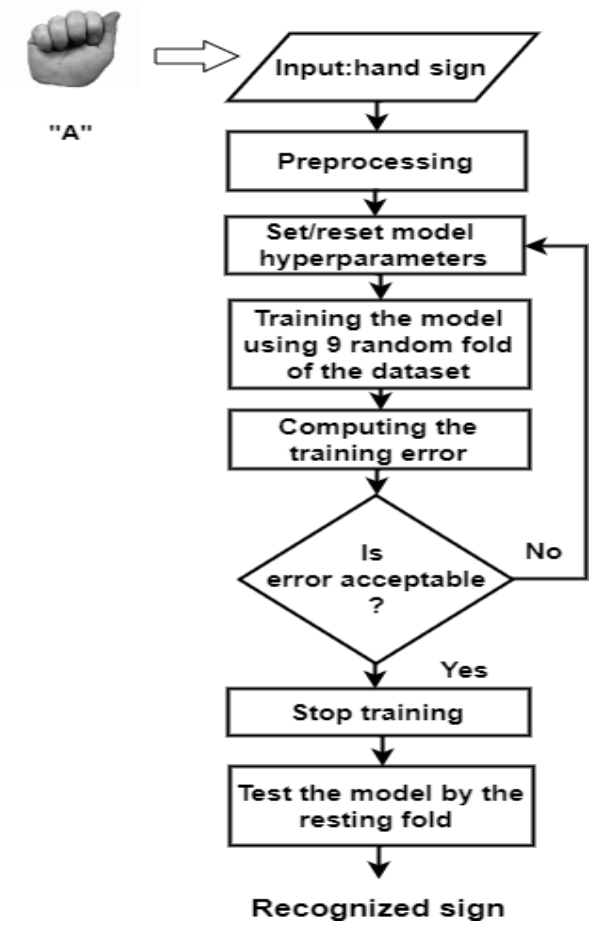

Fig. 1 Methodological steps of the proposed ASL recognition system

The number of kernels used in the convolutional layers may affect the performance of a CNNmodel. There are no standard

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

guidelines for opting the number of kernels in a convolution layer. In this work, we have performed experiment using different number of kernels from 32to 512 at different step size, and finally, the combination which maximize the accuracy was selected. A batch normalization (BN) layer which is responsible for accelerating the trainingprocess and reducing the internal covariate shift is proceeded by some convolution layer. 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 suchusers to build their customized gesture so that the problems faced by persons who aren't ableto talk vocally can be accommodated with technological assistance and the barrier ofexpressing can be overshadowed.

Sign Language Recognition is a very vast topic for research where a lot of work has been done but still various things need to be addressed. The machine learning techniques allow the electronic systems to take decisions based on experience i.e. data. The classification algorithms need two data– training dataset and testing dataset. The training set provides experiences to the classifier and the model is tested using the testing set [6]. Many authors have developed efficient data acquisition and classification methods. The motion data extracted can supply accurate tracking of fingers, hands, and other body parts which leads to robust SLR methodologies development. The vision-based SLR approaches rely on the extraction of discriminative spatial and temporal from RGB images. Most of the vision-based methods initially try to track and extract the hand regions before their classification to gestures . Hand detection is achieved by semantic segmentation and skin colour detection as the skin colour is usually distinguishable easily . Though, because the other body parts like face and arms canbe mistakenly recognized as hands, so, the recent hand detection methods also use the face detection and subtraction, and background subtraction to recognize only the moving parts in a scene.

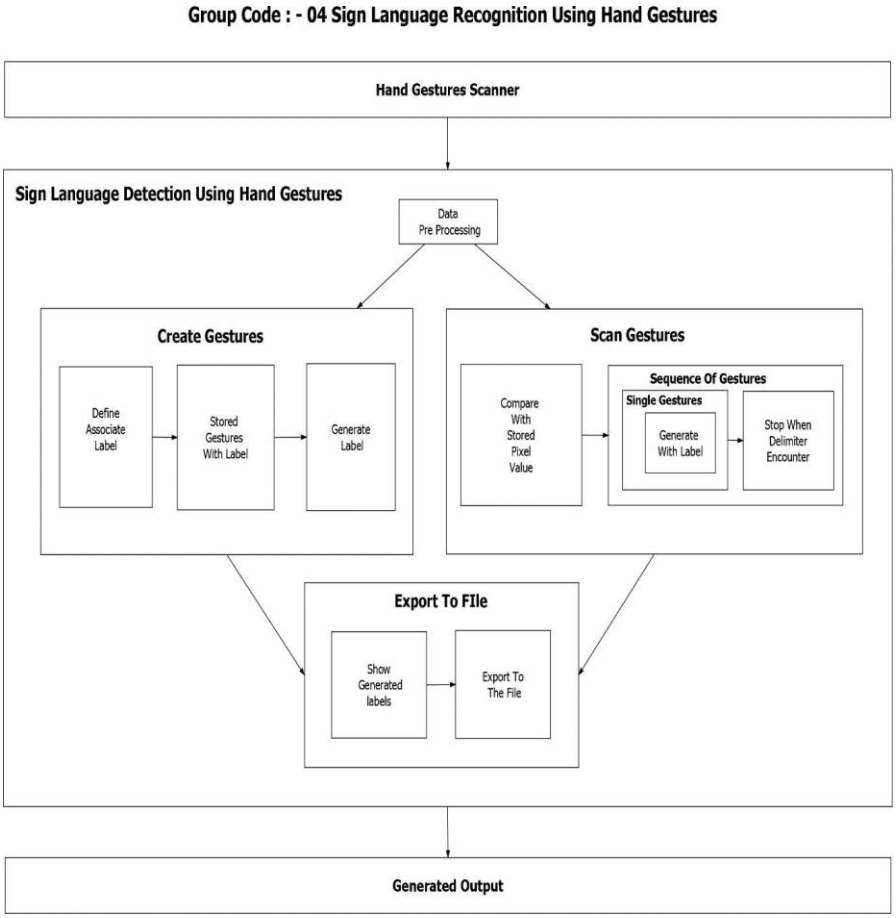

Fig. 2 System Architecture for Sign Language Recognition Using Hand Gestures.

Testing accomplishes a variety of things, but most importantly it measures the quality of the software we are developing. This view presupposes there are defects in the software waiting to be discovered and this view is rarely disproved or even disputed.

Several factors contribute to the importance of making testing a high priority of anysoftwaredevelopment effort. These

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

include:

- Reducing the cost of developing the program.

- Ensuring that the application behaves exactly as we explain to the user for the vast majority of programs, unpredictability is the least desirable consequences of using anapplication.

- Reducing the total cost of ownership. By providing software that looks and behaves as shown in the documentation, the customers require fewer hours of training and lesssupport from product experts.

- Developing customer loyalty and word-of-mouth market share.

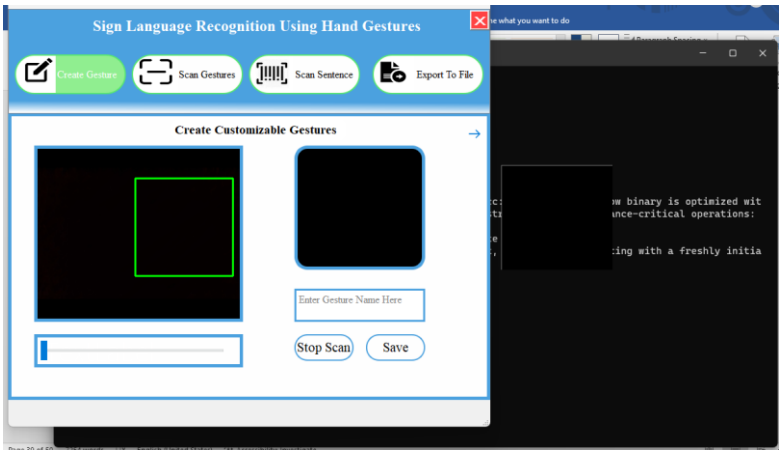

Figure 3: Gesture Recognition Window.

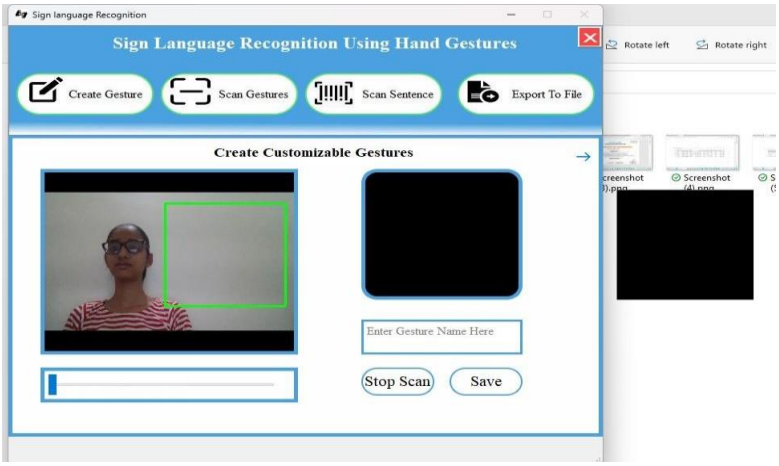

Figure 4: Contact Us Page

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

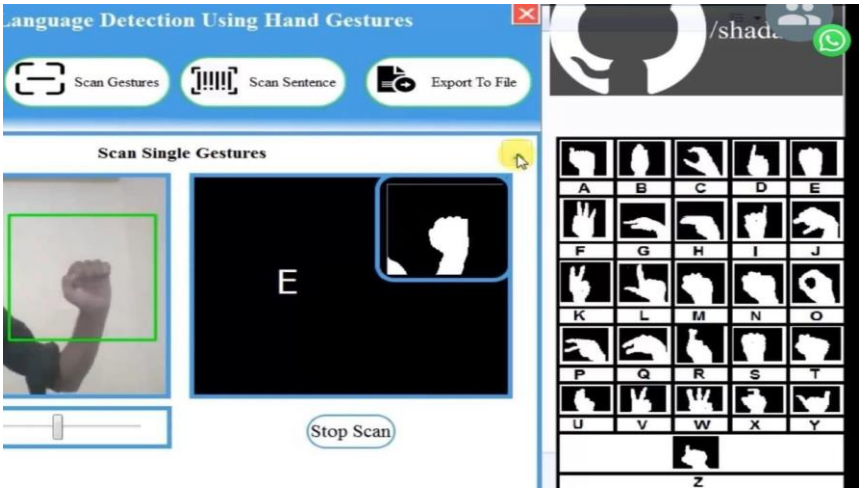

Figure 5: Scanning the sample gesture of alphabet E

IV. CONCLUSION

Upon training the image dataset without any augmentation, the training accuracy achieved was very high (around 99%) but, the real time performance was not up to the mark. It was predicting incorrectly most of the times because in real time hand-gestures were not placed exactly at the center and aligned vertically. In order to overcome this shortcoming, we trained our model by augmenting our dataset. The training accuracy was reduced to 89% but the realtime predictions were predominantly correct. Offline testing of about 9000 augmented images showed an accuracy of 92.7%. From this project/application we have tried to overshadow some of the major problems faced by the disabled persons in terms of talking. We found out the root cause of why they can't express more freely. The result that we got was the other side of the audience are not able to interpret what these persons are trying to say or what is the message that they want to convey. Thereby this application serves the person who wants to learn and talk in sign languages. With this application a person will quickly adapt various gestures and their meaning as per ASL standards. They can quickly learn what alphabet is assigned to which gesture. Add-on to this custom gesture facility is also provided along with sentence formation. A user need not be a literate person if they know the action of the gesture, they can quickly form the gesture and appropriate assigned character will be shown onto the screen. Concerning to the implementation, we have used TensorFlow framework, with keras API. And for the user feasibility complete front-end is designed using PyQT5. Appropriate userfriendly messages are prompted as per the user actions along with what gesture means which character window. Additionally, an export to file module is also provided with TTS(TextTo-Speech) assistance meaning whatever the sentence was formed a user will be able to listen to it and then quickly export along with observing what gesture he/she made during the sentence formation.

FUTURE WORK

➢ It can be integrated with various search engines and texting application such as google, WhatsApp. So that even the illiterate people could be able to chat with other persons, or query something from web just with the help of gesture.

➢ This project is working on image currently, further development can lead to detecting the motion of video sequence and assigning it to a meaningful sentence with TTS assistance.

➢ We can develop a model for ISL word and sentence level recognition. This will require a system that can detect changes with respect to the temporal space.

➢ We can develop a complete product that will help the speech and hearing-impaired people, and thereby reduce the communication gap.

ACKNOWLEDGMENT

We would like to thank **Mrs. Khateeja Ambareen (**Dept. of CSE, GSSSIETW, Mysuru) for their valuable guidance, feedback and encouragement to complete the project.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10574

REFERENCES

1. 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-anddumb-people-while using-todays-common-tech-like-phones-and-PCs, [Accessed April 06, 2019].
2. NIDCD, "american sign language", 2017 [Online]. Available: https://www.nidcd.nih.gov/health/american-sign-language, [Accessed April 06, 2019].
3. Suharjito MT, "Sign Language Recognition Application Systems for Deaf-Mute People A Review Based on Input-Process-Output", 2017 [Online]. Available:
4. https://www.academia.edu/35314119/Sign_Language_Recognition_Application_Systems_for_Deaf-Mute_People_A_Review_Based_on_Input-Process-Output [Accessed April 06, 2019].
5. M. Ibrahim, "Sign Language Translation via Image Processing", [Online]. Available: https://www.kics.edu.pk/project/startup/203 [Accessed April 06, 2019].
6. NAD, "American sign language-community and culture frequently asked questions", 2017 [Online]. Available: https://www.nad.org/resources/americansign-language/community-and-culturefrequently-asked-questions/ [Accessed April 06, 2019].
7. Sanil Jain and K.V.Sameer Raja, "Indian Sign Language Character Recognition" , [Online]. Available: https://cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf [Accessed April 06, 2019].