International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

RECOGNITION OF SIGN LANGUAGE IN REAL TIME

Ais Athania, Student, MCT's RGIT Mumbai India, awaisathania@gmail.com

Krishna Sanjay Gupta, Student, MCT's RGIT Mumbai India, guptakrishna613@gmail.com

Karima Khan, Student, MCT's RGIT Mumbai India, kareemakhan95@gmail.com

A.E. Patil, Professor, MCT's RGIT Mumbai India, abhay.patil@mctrgit.ac.in

Abstract: Failure to communicate is regarded as a weakness. Individuals with this incapacity utilize various modes to speak with others, there are a number of strategies accessible for their correspondence; one such normal technique for correspondence is gesture-based communication. Creating communication via gestures application for hard of hearing individuals can be significant, as they'll have the option to discuss effectively with even the individuals who don't comprehend communication via gestures. Our undertaking targets making the fundamental stride in spanning the correspondence hole between ordinary individuals, hard of hearing and unable to speak individuals utilizing sign language. The main focus of this work is to make a vision based system to spot signing gestures. The reason for selecting a system supported vision relates to the very fact that it provides an easier and more intuitive way of communication between a person's and a computer. Despite much previous works, traditional vision based hand gesture recognition methods are still far away from satisfactory for several real world applications. In this work, we propose a deep learning-based methodology for building an American Sign Language (ASL) for static alphabet recognition system. Here, we review various existing methods in sign language recognition and implement Convolutional Neural Network (CNN) architecture for American Sign Language (ASL) static alphabet recognition.

Keywords —American Sign Language, Classification, Convolutional Neural Network, Feature extraction, Open CV, Pre processing, Vision based system

I. INTRODUCTION

Deafness and vocal impairment are some of the major disabilities faced by human beings for centuries. This problem hinders an individual from speaking in verbal languages to the outside world resulting in isolation from the remainder of the main verbally communicating society. They use sign language for communication with people but this is limited to them or some relatives. A manual sign language interpreter is not necessarily a feasible option, because it often interferes with the person's privacy. So a solution to this problem is an automated sign language translator which can translate sign language into natural language.

Here for gesture recognition, we are using image processing and computer vision. Gesture recognition enables computers to know human actions and also acts as an interpreter between computer and human. This could provide the potential to humans to interact naturally with computers with no physical contact with mechanical devices. Gestures are performed by deaf and dumb community to perform sign language when broadcasting audio is impossible, or typing and writing is difficult, but there is the vision possibility. At that point sign language is the only way for exchanging information between people. Generally, sign language is used by everyone when they do not want to speak, but this is the only way of communication for deaf and dumb community. Sign language is additionally serving an equivalent meaning as speech does. This is employed by deaf and dumb community everywhere in the world but in their regional form like ISL, ASL. Sign language is often performed by using Hand gestures either by one hand or two hands.

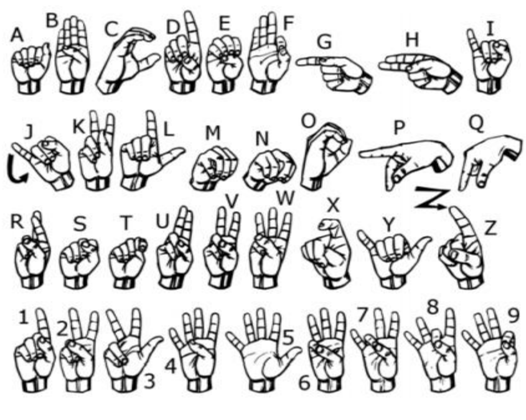

Fig1.Finger Spelling American Sign Language

II. LITERATURE SURVEY

The paper [1] provides an overview of Indian sign language, its dialects and variants, and current attempts to

 DOI : 10.35291/2454-9150.2021.0130

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

standardize it. In this to acknowledge the static single hand gestures of a subset of Indian sign language, a proposed methodology is discussed. The goal is to make a system which will identify gestures of the human hand and use them to convey information without the utilization of an interpreter. The study Indian sign language and its varieties has been summarized in this paper Also a simple recognition system is proposed. The proposed recognition system aims for signer-independent operation and utilizes one web camera for data acquisition to make sure user-friendliness.

In [2] the authors have classified 140 classes which include finger-spelling numbers, alphabets, and other common phrases as well. For the dataset, a Kinect sensor is used. RGB images of resolution 640 x 480 are captured along with their depth data. Depth values of each pixel are captured in the depth data of the image. Each pixel of the image corresponds to a value in the depth data file. The system proposes the tactic for a completely unique low-cost and easy-to-use application for the Microsoft Kinect camera, ie Indian sign language recognition.

In [3] the authors proposed a system that considers 24 alphabets of ISL, each alphabet consists 10 samples thus a total of 240 images existed in the database. They used only five significant Eigen vectors out of fifty because all other Eigen vectors are very small, and that they neglected the undesired Eigen vectors. This provides advantages like compression of knowledge, dimensionality reduction without much loss of data, dropping the first variables into a lower number of orthogonal or non-correlated synthesized variables. With this new technique, recognition rate is 97.00%.

III. METHODOLOGY

To recognize the signs represented by each of these images, we trained a Convolutional Neural Network (CNN). The CNN architecture includes 3 groups of 2 convolution layers, a max pool layer and a dropout layer, and 2 groups of fully connected layers followed by a dropout layer and one final output layer. In addition, frames are captured using the webcam's live feed and OpenCV libraries are used for preprocessing.

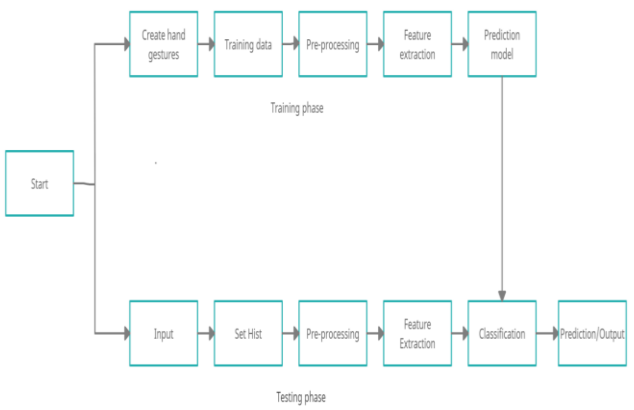

Fig 2.System flow diagram

A. Image acquisition

The initial step of Image Acquisition is of acquiring a picture during runtime through the integrated camera and while acquiring these images are going to be stored within the directory after they're captured and after capturing the recent image it will be acquired and that image will be compared with images stored for that letter in the database using the SIFT algorithm. The images are going to be captured through the basic code of opening a webcam through OPENCV then capturing the image through frames per second which will be stored in another directory where all the input images are stored in another directory and the newly captured image is picked up and compared with the given set of images.

B. Pre-processing

The primary goal of pre-processing is to improve image data by enhancing image attributes or reducing unnecessary variations for subsequent processing. Preprocessing is additionally referred to as an effort to capture the important patterns which express the individuality in data without noise or unwanted data which incorporates cropping, resizing, and grey scaling.

C. Feature Extraction

In any object, there are many elements, interesting points in an object, which can be extracted to give a description of the "element" of the object. SIFT image features provide plenty of object features that are unaffected by many of the problems encountered in other approaches, such as object scaling and rotation. The SIFT method, for image processing, takes an image and converts it into a "large collection of local feature vectors". Each feature vector does not change to any scale, rotation, or image translation. Different layers are been used in the Convolutional neural network to extract low-level and high-level features of the images such as convolution layer, max-pooling layer, and fully connected layer.

D. Classification

Using the Soft-max Classification process, the model is able to distinguish between domination and certain low-level features in images and isolation. Soft-max offers decimal probabilities for each class in a multi-class problem. Just before the output layer, the Soft-max is applied through a neural network layer.. The Soft-max layer should have the same number of nodes as the output layer.

E. Text to Speech

When the character is selected, recognized text is converted to speech based on the recognized signs using speech conversion.

 DOI : 10.35291/2454-9150.2021.0130

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

IV. Result

The system was checked with various alphabets and signs, and the system's prediction is fairly reliable. Figures 3, 4, and 5 depict the prediction of numbers, alphabets, and gestures, respectively.

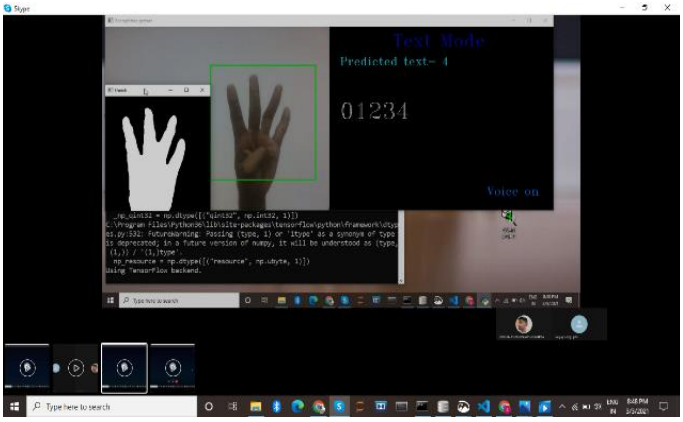

Fig 3.Prediction of numbers

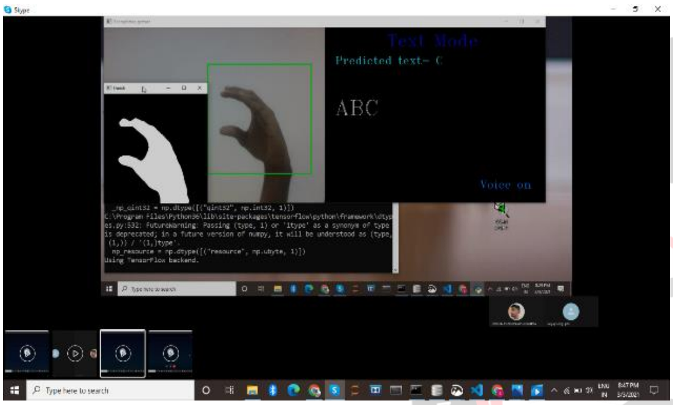

Fig 4.Prediction of alphabets

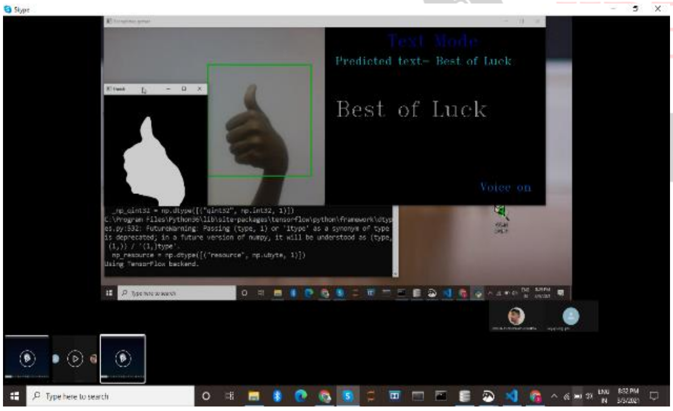

Fig 5.Prediction of word 'Best of luck'

Our project's major flexibility is that we can create our own gestures and train the model accordingly (Fig 6). We kept the 1200 scale. It means that it will capture 1200 pictures of the specific gesture and store them in the folder we have created (Fig 7 and 8). Also, the database will be updated accordingly (Fig 9).

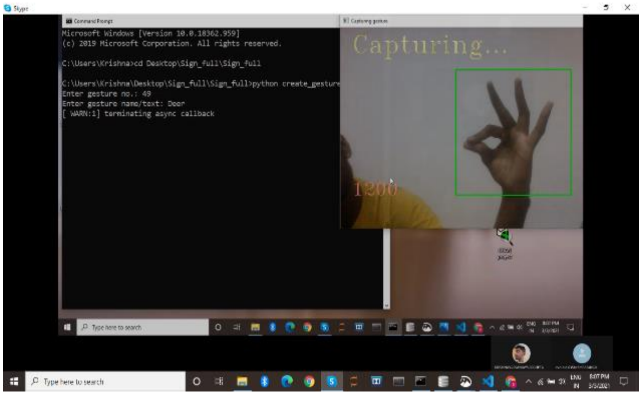

Fig 6.Creating gestures

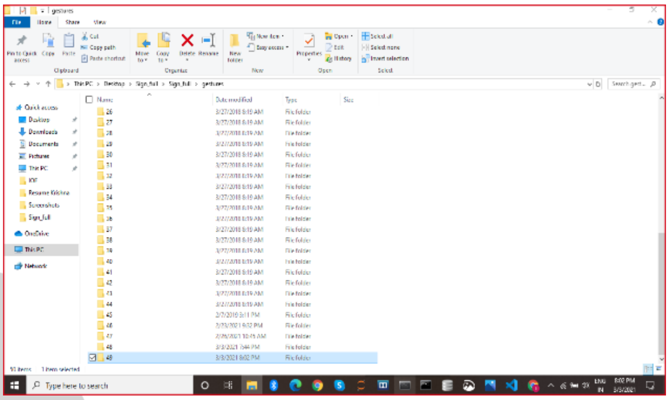

Fig 7.creating a new folder for new gesture

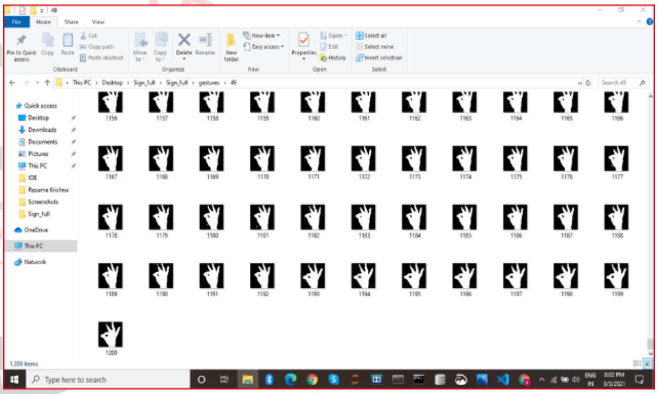

Fig 8.training data of newly created gesture

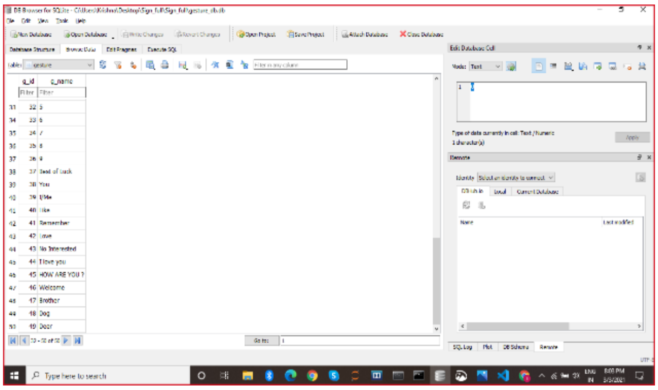

Fig 9.Snapshot of the database after adding new gesture

V. Conclusion

Many people who are unfamiliar with sign language find it difficult to communicate without the assistance of an

DOI : 10.35291/2454-9150.2021.0130

ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

interpreter. This work would assist in the achievement of high performance in sign language recognition. Furthermore, it transcribes sign language symbols into plain text and converts the interpreted sign language into speech, allowing real-time communication. Because of the system's flexibility, additional gestures can be added to make it more customized for usage. This system accepts static hand gestures through a webcam as input and uses OpenCV libraries for image acquisition and preprocessing. For feature extraction Convolutional neural network is used and the soft-max classification technique recognizes the signs. Also, the environment plays an important role in sign recognition; proper hand positioning, improved lighting, and so on improve the recognition rate. Our system uses a histogram for input image enhancement, thus increasing the recognition rate. As a result, this system acts as a link between deaf and mute people and the general public.

VI. Future Work

While sign language involves manual, facial, and body expressions, this work focuses solely on finger signs. As a result, this system can be used as a module for more advanced work. Furthermore, the database used in this work is much smaller, but since the database is flexible, several more terms can be added for better prediction.

References

[1] Pravin R Futane, Rajiv V. Dharaskar, “HASTA MUDRA - An interpretation of Indian sign hand gestures,”2011 3rd International Conference on Electronics Computer Technology (ICECT 2011)

[2] Zafar Ahmed Ansari and Gaurav Harit, “Nearest neighbour classification of Indian sign language gesturesusing kinect camera,” Sadhana Vol. 41, No. 2, February 201 , pp. 161–182

[3] Joyeeta Singha, Karen Das “Recognition of Indian Sign Language in Live Video,” International Journal of Computer Applications (0975 – 8887), vol. 70– No.19, May 2013

[4] A. Das, S. Gawde, K. Suratwala and D. Kalbande, "Sign Language Recognition Using Deep Learning on Custom Processed Static Gesture Images," 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, India, 2018, pp. 1-6

[5] D. Deora and N. Bajaj, "Indian sign language recognition," 2012 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking, Surat, India, 2012, pp. 1-5

[6] S. C.J. and L. A., "Signet: A Deep Learning based Indian Sign Language Recognition System," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0596-0600

[7] L. Zheng, B. Liang and A. Jiang, "Recent Advances of Deep Learning for Sign Language Recognition," 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, NSW, Australia, 2017, pp. 1-7

[8] M. Hasan, T. H. Sajib and M. Dey, "A machine learning based approach for the detection and recognition of Bangla sign language," 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec), Dhaka, Bangladesh, 2016, pp. 1-5

[9] A. S. Nikam and A. G. Ambekar, "Sign language recognition using image based hand gesture recognition techniques," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, India, 2016, pp. 1-5

[10] Sahoo, Ashok & Mishra, Gouri & Ravulakollu, Kiran “Sign language recognition: State of the art,” 2014 ARPN Journal of Engineering and Applied Sciences. 9. 116-134.

 DOI : 10.35291/2454-9150.2021.0130