INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 04 | APRIL - 2023 IMPACT FACTOR: 8.176 ISSN: 2582-3930

Detection and Conversion of Dactylology to Speech

G L Naga Hari Prasath
Department of Computing Technologies,
SRM Institute of Science and Technology,
Chennai, India.
gn1729@srmist.edu.in

Naveen A
Department of Computing Technologies,
SRM Institute of Science and Technology,
Chennai, India.
na7277@srmist.edu.in

SREEKUMAR K
Department of Computing Technologies,
SRM Institute of Science and Technology,
Chennai, India.
sreekumk@srmist.edu.in

Abstract— Hand gestures are one of the nonverbal communication modalities used in sign language. When deaf or dumb persons have hearing or speech problems, they use it to communicate with others or among themselves. Regular people can connect and communicate with one another with ease, but those with hearing and speaking impairments have difficulty doing so without a translator. Moreover, it can be difficult for them to converse with people who are blind. Various sign language systems have been developed around the world, but they are neither adaptable nor cost-effective for end users. As a result, software that can automatically recognize sign language is shown as a proposed method in this study in order to enable deaf and dumb individuals to communicate with normal people and each other more efficiently. In addition, our approach facilitates communication between blind and deaf individuals by translating identified sign language into voice using NLP techniques.

I. INTRODUCTION

Humans depend heavily on communication since it allows us to express ourselves. Humans can communicate using speech, body language, reading, writing, or visual assistance, with speaking being one of the most popular methods. A communication gap does, regrettably, exist for the minority who are speech and hearing impaired. To communicate with them, visual aids or an interpreter are used. These techniques, nevertheless, are highly time-consuming and expensive, and they cannot be employed in an emergency. For the most part, sign language involves manual communication to communicate meaning. In order to express one's ideas, one must combine the shapes, orientations, and movement of the hands, arms, and body at the same time.

Word level association, which uses hand movements to express word meaning, and fingerspelling, which spells out words character by character, are both components of sign language. Finger spelling is an important technique in sign language for transmitting names, addresses, and other words that lack value at the word level. Despite this, fingerspelling is not widely used since it is difficult to learn and use. However, no universal sign language exists, and even fewer people are familiar with it, making it a poor substitute for spoken communication. Hence, a system that can automatically recognize sign language motions and translate them into voice is required. By implementing such a system, the social divide between deaf, blind, and normal people would be reduced.

II. LITERATURE SURVEY

There is a lot of research efforts in the field of developing recognition systems for sign language around the world.

Zabulisy et al. [1] proposed a vision-based hand gesture recognition system for Human-Computer Interaction. In vision-based approaches introduced to overcome these problems, Mohandes[2-5] introducedaprototypesystemto recognize the Arabic sign language based on Support Vector Machine (SVM) and also an automatic translation system to translate Arabic text to Arabic sign language.

AlJarrah and Halawani [6] developed a neuro-fuzzy system that deals with images of bare hand signs and achieved a recognition rate of 93.55%. In [7], Al-Rousan and Hussain built an adaptive neuro-fuzzy interference system for letter recognition. A colored glove was used to ease the process of segmenting the hands region. e rec ognition accuracy achieved was 95.5%.

Automatic sign language recognition approaches often combine ideas from computer vision and speech recog nition. A variety of sign language-specific visual pre processing and features have been proposed in prior work, including ones based on estimated position and movement of body parts (e.g. hand, head) combined with appear ance descriptors [8, 9]. Recent work has had success with convolutional neural network (CNN)-based features [10,11,12,13,14,15,16,] Much previous work on sign lan guage recognition, and the vast majority of previous work on fingerspelling recognition, uses some form of hand de tection or segmentation to localize the region(s) of interest as an initial step. Kim et al. [18, 19, 17] estimate a signer dependent skin color model using manually annotated hand regions for fingerspelling recognition. Huang et al. [15] learn a hand detector based on Faster R-CNN [33] using 2There has also been work on sign language recognition using other modalities such as depth sensors (e.g., [32, 15]). Here we consider video only input, as it is more abundant in naturally occurring online data. manually annotated signing hand bounding boxes, and ap ply it to general sign language recognition. Some sign language recognition approaches use no hand or pose pre processing as a separate step (e.g., [22]), and indeed many signs involve large motions that do not require fine-grained gesture understanding. However, for fingerspelling recognition it is particularly important to understand fine-grained distinctions in handshape.

III. PROPOSED WORK

In this proposed system the dataset is given by us by taking 100 approx. pictures of each sign for each alphabet for training the model, and image to be detected need not be processed digitally which consumes more time it's done using Hand tracking modules and classification modules. Not only these, in our proposed system the sign detected is being converted into speech so that it also helps blind people for communication in ease.

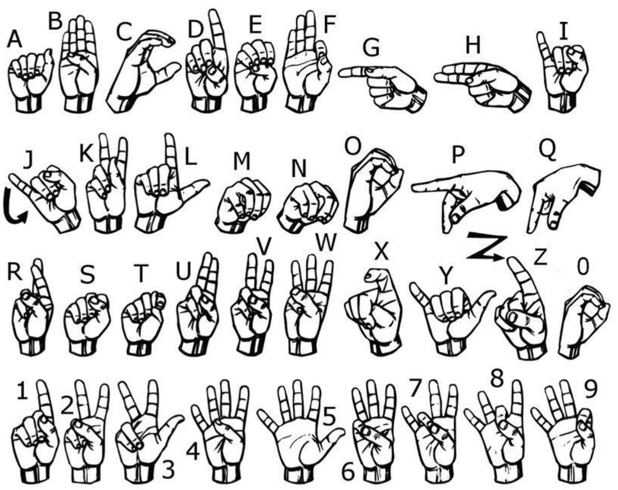

Fig 1 : American Sign Language Representation

IV. IMPLEMENTATION

DATA COLLECTION:

For this sign data, we have not taken predefined dataset, we manually collect the data by replicating each sign form alphabet a to z in front of the camera using cz module, for each sign we take approx.100 pictures in different angle and in different direction to increase the accuracy of detection.

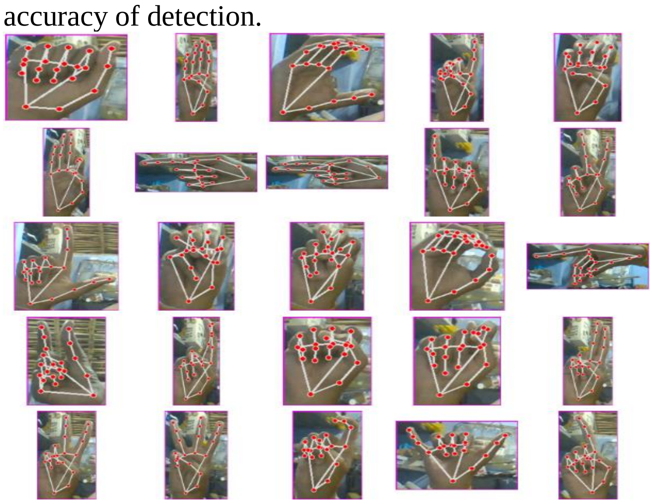

Fig 2. Sign Representation Data

MODEL TRAINING:

Once the data set is ready, grouping of the data is done accordingly and the trained model is extracted in keras module format. This module is used in the next phase where the sign detection takes place.

VOLUME: 07 ISSUE: 04 | APRIL - 2023 IMPACT FACTOR: 8.176 ISSN: 2582-3930

SIGN DETECTION:

The sign showed in camera is processed into frames then each frame is processed to accumulate them in the same size in pixels, doing this way increases the speed of detection of the sign if not done it takes some time to convert the frame to a particular size then the detection takes place. The sign is recognized and classified with referring to the pre trained model which we feed.

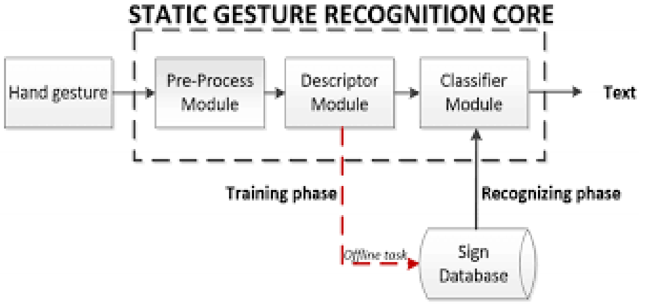

Fig.3 Recognition Flow

TEXT TO SPEECH CONVERSION:

The detected sign is stored as text and is converted to speech. Pyttsx3 is the text-to-speech library used in this conversion. This pyttsx3 works in offline and is more compatible. The text is broken into tokens and uses tokenizersto translate.

V. OUTPUT

Input is the hand gesture shown by the user in camera, the following outputs are which we got.

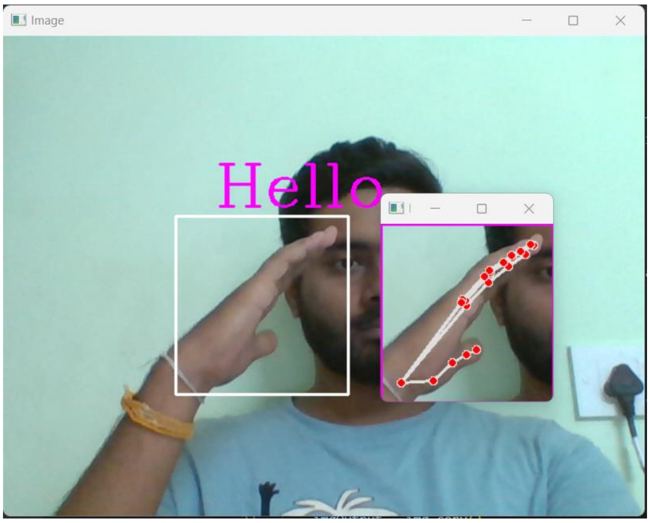

Fig.4 'Hello' as Output

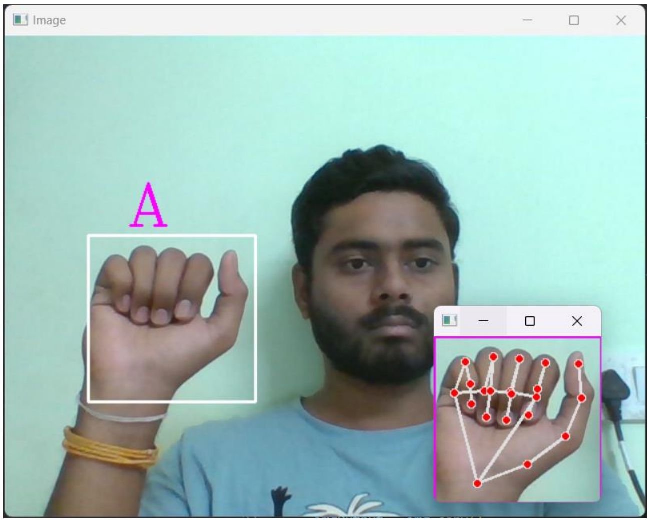

Fig.5 'A' as output

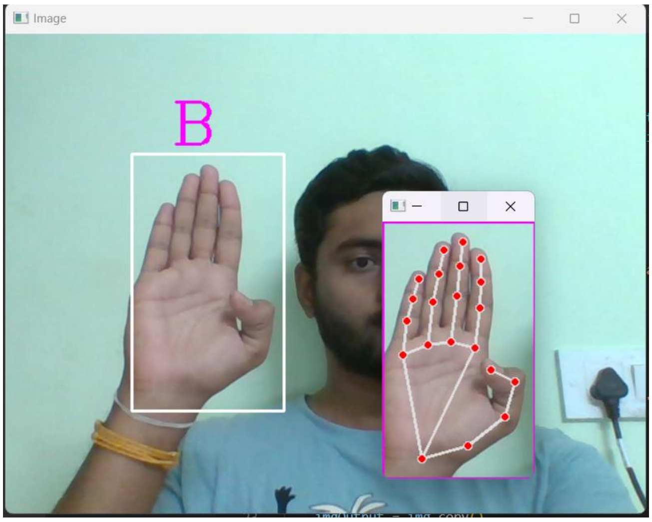

Fig.6 'B' as output

VI. CONCLUSION

Various Test cases have been conducted by taking different alphabets and even words using sign language and the system was able to distinguish between the alphabets and was able to identify each and every alphabets and was capable of converting them into speech , we even tried with the combination of alphabets to form words in sign language even then our system was able to identify each letters of the word and converted into speech, i.e. sign language to text then to speech.

REFERENCES

[1] X. Zabulis, H. Baltzakis, and A. A. Argyros, "Vision-based hand gesture recognition for human-computer interaction," e universal access handbook, vol. 34, p. 30, 2009.

[2]M.Mohandes,"AutomatictranslationofArabictexttoArabic sign language," AIML Journal, vol. 6, no. 4, pp. 15–19, 2006.

[3] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign lan guage recognition using the leap motion controller," in Proceedings of the 2014 IEEE 23rdInternational Symposium on Industrial Electronics (ISIE), pp. 960–965, IEEE, Istanbul, Turkey, June 2014, þ.

[4] M. Mohandes, M. Deriche, U. Johar, and S. Ilyas, "A signer independent Arabic sign language recognition system using face detection, geometric features, and a hidden markov model," Computers & Electrical Engineering, vol. 38, no. 2, pp. 422–433, 2012, þ.

[5] M. Mohandes, "Arabic sign language recognition," in Pro ceedings of the International Conference of Imaging Science, Systems, and Technology, pp. 753–759þ, Las Vegas, NV, USA, 2001.

[6] O. Al-Jarrah and A. Halawani, "Recognition of gestures in Arabic sign language using neuro-fuzzy systems," Artificial Intelligence, vol. 133, no. 1-2, pp. 117–138, 2001.

[7] M. Al-Rousan and M. Hussain, "Automatic recognition of Arabic sign language finger spelling," International Journal of Computers and eir Applications, vol. 8, pp. 80–88, 2001.

[8] Richard Bowden, David Windridge, Timor Kadir, Andrew Zisserman, and Michael Brady. A linguistic feature vector for visual representation of sign language. In ECCV, 2004.

[9] Mahmoud M. Zaki and Samir I. Shaheen. Sign language recogntion using a combination of new vision based features. In Pattern Recognition Letters, 2011.

[10] Oscar Koller, Hermann Ney, and Richard. Bowden. Deep hand: How to train a CNN on 1 million hand images when your data is continuous and weakly labelled. In CVPR, 2016.

[11] Oscar Koller, Sepehr Zargaran, and Hermann Ney. Re-sign: Re-aligned end-to-end sequence modelling with deep recur rent CNN-HMMs. In CVPR, 2017.

[12] Oscar Koller, Sepehr Zargaran, Ralf Schlter, and Richard Bowden. Deepsign: HybridCNN-HMMforcontinuoussign language recognition. In BMVC, 2016.

[13] Oscar Koller, Sepehr Zargaran, Hermann Ney, and Richard Bowden. Deep sign: Enabling robust statistical continuous sign language recognition via hybrid CNN-HMMs. Inter national Journal of Computer Vision, 126(12):1311–1325, 2018

[14] Runpeng Cui, Hu Liu, and Changshui Zhang. Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. In CVPR, 2017.

[15] Bowen Shi and Karen Livescu. Multitask training with unla beled data for end-to-end sign language fingerspelling recog nition. In ASRU, 2017.

[16] Bowen Shi, Aurora Martinez Del Rio, Jonathan Keane, Jonathan Michaux, DianeBrentari, Greg Shakhnarovich, and Karen Livescu. American Sign Language fingerspelling recognition in the wild. In SLT, 2018