Efficient Gesture based Language Recognition using SVM and Lloyd's Algorithm

Abdul Khader, Muhammad Thouseef, Akbar Ali, Ahamad Irfan

ABSTRACT--- undertaking in the field of human-computer interaction (HCI) and computer vision. 10 years prior to the undertaking appeared to be practically unsolvable with the data given by a single RGB camera. In this work, we have actualized a presumable exact strategy to perceive static gestures or image frames from a live camera or video data. As Hand Gesture Recognition is identified with two noteworthy fields of image processing and AI (machine learning), in this way, this report likewise refers to the different tools and APIs that can be utilized to implement different strategies and methods in these fields.

Keywords — Sign language recognition, live camera, SVM, Gesturing, Lloyd's algorithm, audio output, TensorFlow, CNN

I. INTRODUCTION

Deaf-mute individuals need to speak or communicate with typical individuals for their day by day routine and schedule. The deaf and dumb individuals all through the world utilize the sign language or the gesture-based communication to speak with other people. Be that as it may, it is conceivable just for the people who have experienced special preparing and training to comprehend the language. Communication via gestures uses hand motions and different methods for non-verbal practices to pass on their proposed meaning and intended word strength. It includes consolidating hand shapes, direction and hand movements, arms or body movement, and outward appearances(facial) at the same time, to smoothly express the speaker's contemplations.

A gesture is an example which might be static, dynamic or both [1], and is a type of interchanges which are non-verbal in which substantial movements pass on information. Motions incorporate movement of head, hands, fingers or other body parts [2]. Regardless, Gesture Recognition, by and large, suggests the whole strategy of following human movements, to their depiction and change to semantically critical directions. Gesture Recognition and all the more explicitly hand motion acknowledgment can be utilized to upgrade Human-Computer Interaction (HCI) and improve the viable usage of the accessible data stream.

Revised Manuscript Received on July 10, 2019.

Abdul Khader, Assistant Professor, Computer Science & Engineering Department, Bearys Institute of Technology Innoli, Mangaluru.Karnataka, India (E-mail: Khader.mtech14@gmail.com)

Muhammad Thouseef, Computer Science & Engineering Department, Bearys Institute of Technology Innoli, Mangaluru. Karnataka, India (E-mail: muhammadthouseef05@gmail.com)

Akbar Ali, Computer Science & Engineering Department, Bearys Institute of Technology Innoli, Mangaluru. Karnataka, India (E-mail: akbardestiny07@gmail.com)

Ahamad Irfan, Computer Science & Engineering Department, Bearys Institute of Technology Innoli, Mangaluru. Karnataka, India (E-mail: ahamad.irfa@gmail.com)

Communication via gestures and its recognition remains an extensive issue because of the intricacy of visual analysis and quickness of basic changes in marked gestures or motions [3]. Most of the countries have their own standardized sign languages (ISL-Indian Sign Language). Segmentation is the primary challenge in any form of gesture recognition as a method. The continuous or moving object captured image has many challenges. Several segmentation algorithms show high-efficiency rates for a particular database of gestures; but, they're quickly rendered less effective as video backgrounds degrade, or because the personal hand positions deviate from the middle of captured frame space. There are many more technologies projected and proposed as an RGB-based image segmentation technique conjointly like YCbCr- based mostly image segmentation technique, none of those systems provides a correct means for real-time human-computer interaction where by it gets a great deal and a lot more difficult to predict the output using mentioned methods.

The second challenge within the gesture recognition is separating sign acts or gesture forms from the transitional movements between the visual gestures [4]. On the off likelihood that the speed of motions shifts too, the presence of image edges may well be too flighty to even take into account producing nice scientific or analytical assessments and estimates. If the speed of gestures varies, then the looks of image frames conjointly capricious analyse.

II. EXISTING SYSTEM

There are several methods of implementation and systems existing at present. Every one of them is diverse in a few or different capacities. Some have distinctive calculation and methods utilized for perceiving the hand motion [5]. Hand gesture acknowledgment framework is considered as a path for increasingly natural and capable human-computer interaction tool.

Majority of the existing systems are based on python programming implemented by using Machine Learning and Deep learning. Some of the systems are built and developed using MATLAB programming where computing and sigmoid function thresholds are clearly shown using the built-in functions of the software which helps to ease the modification and debugging.

III. PROPOSED SYSTEM

We aim to make the sign language recognition system to exhibit a proficient and precise mechanism to translate content or discourse, subsequently the "dialog communication" between the hard of hearing and hearing person will be smooth. There is no institutionalized communication via gestures for every single hard of hearing individuals over the world. In any case, gesture-based communications are not widespread, likewise, with spoken dialects, these vary from area to locale. An individual who can talk and hear appropriately (ordinary person) can't speak with the hard of hearing and moronic individual except if he/she knows about gesture-based communication. A similar case is relevant when a hard of hearing and dumb individual needs to speak with an ordinary person or visually impaired person. There might be a lot of communication assistants in the current era. But still, real-time systems with the help of external devices such as camera and audio will be the basic level implementations compared to large scale modern technologies.

A. Advantages of the Proposed System

From this EFFICIENT GESTURE BASED LANGUAGE RECOGNITION application, the deaf and dumb can easily get the alphabets printed on the screen with the help of Sign language shown by them. If this is implemented online, then people if access this application, they can remotely avail the outputs (alphabets).

Highlights:

- Real time Webcam Input
- Video/Image Input
- Simultaneous Output (Alphabets)
- Text to Audio output

IV. DESIGN AND IMPLEMENTATION

The application is designed and implemented in a systematic way so that we could get our predicted result as output without causing any limitations using the set of APIs and Algorithms.

A. Work Flow of the Application

This project will show the real-time output as alphabets when the trained hand signs are shown in front of the camera. Except for a few letters with a dark background or different hand shapes, rest all alphabets works well during prediction. The convolutional Neural Network (CNN) operation is implemented is this concept for the purpose of edge detection and formation of image matrix with the help of TensorFlow.

The layer-based matrix filtering and the size of matching portions in the decided area in the frame are basically the CNN. The convolutional, pooling and the fully connected layer architecture helped to focus the areas in the frame. And also, the resulting alphabets are then converted as an audio output using the internet to acquire the audio file.

Full implementation is shown in Fig. 1.

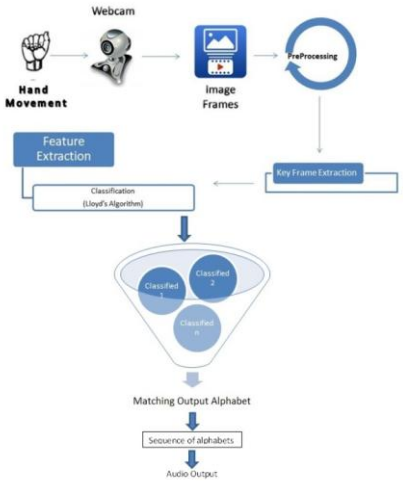

Fig. 1 Sign Language detection system

B. Support Vector Machine (SVM)

This is a supervised machine learning algorithm that can be utilized for both regression or classification challenges. In any case, it is for the most part utilized in classification problems. In this algorithm, we plot every data item as a point in n-dimensional plotting where n is the quantity of highlights with the value of every component being the value of the specific coordinate. At that point, we perform grouping (classification) by finding the hyperplane that isolates and separates the two classes (take a look at the underneath preview).

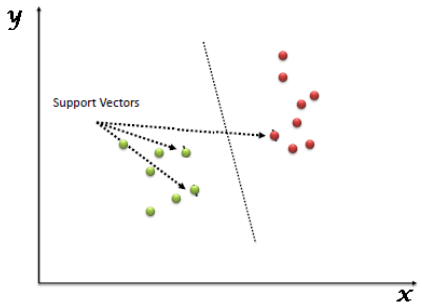

Fig. 2 Support Vector Machine Hyperplane

C. Lloyd's Algorithm

Lloyd's algorithm, otherwise called Voronoi iteration or relaxation, is an algorithm named after Stuart P. Lloyd for finding equally divided arrangements of points in subsets of Euclidean spaces and segments of those subsets into all around moulded and consistently estimated arched cells. Like the firmly related k-means clustering algorithm, it over and again finds the centroid of each set in the partition and after that re-segment the input as indicated by which of these centroids is nearest. In any case, Lloyd's algorithm contrasts from k-means clustering in that its information is a constant geometric locale as opposed to a discrete arrangement of points. Therefore, when re-apportioning the input, Lloyd's algorithm utilizes Voronoi charts instead of just deciding the closest point to every one of a limited arrangement of points as the k-means algorithm does.

Step 1: Start
Step 2: The initial image is taken for classification
Step 3: The image is changed into a pixel matrix
Step 4: Using Lloyd Iteration, clusters are made based on centroids
Step 5: Compute Distortion
Step 6: Perform the test operation and check Validation of Test cases
Step 7: If step 6 is Successful then continue to next step, else go to Step 2 and repeat the procedure until the test case becomes valid.
Step 8: End the process.

B. Sign Language Recognising

The real-time detection and recognition are done when the hand sign language is shown in front of the camera and simultaneously output is obtained, as shown in Fig. 3.

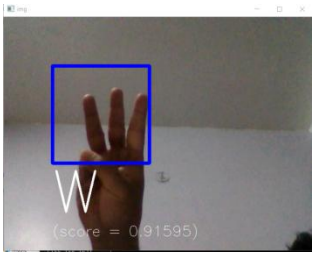

Fig. 3 Recognized Alphabets with its accuracy

C. Audio Output

After the text recognition is done the audio sequence output is played, as shown in Fig. 4.

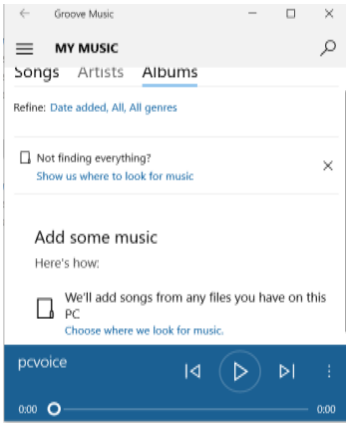

Fig. 4 Sequence Output is played as Audio

V. RESULT ANALYSIS

A. Real Life Testing

To test the system for robustness in identifying, detecting, tracking and following diverse skin hues in various lighting condition, with various backgrounds and shadows, and after that perceiving the different gestures we completed a different real-life test as shown in figure 5. We additionally played out a few gestures of each class (alphabets) in two diverse lighting condition, with various background and with 2 unique subjects anddetermined the accuracy of classification.

Fig. 5 Real time testing

B. Handshape Classifiers

There are various types of handshape classifiers or sign languages present in the world. Among all, American Sign Language (ASL) is one most widely used handshape classifier[6]. We use these sign languages in our project for creating the database and to recognize the alphabets. We had to modify a few letters of ASL as shown in fig 6 to predict our project results more accurately.

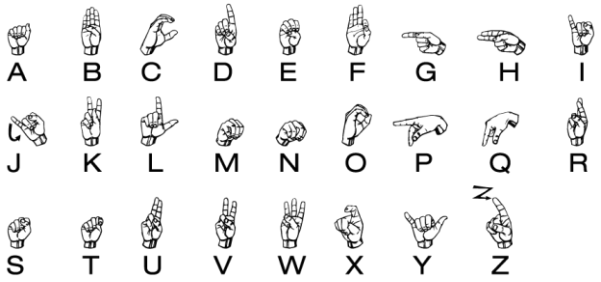

Fig. 6 ASL Handshape Classifiers

C. Test Analysis

The different tests carried out were of using different hands and most of them were having the white-coloured background. We have got the successive result of test phase done by same background and same hand while training the model, and their result was 100% success, unlike in others it was unable to recognize, or the accuracy was very less because some hands and background caused complexity due to the light, colour mismatching and darkness.

Below listed in Table I and Table II are the experimental analysis data based on ten different tests and their success rate is shown.

Table 1 Part 1 of Analysis Table

Alpha-bets	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10	Success Rate (%)
A	True	True	True	True	True	False	True	True	False	True	80
B	True	True	False	True	True	False	True	True	False	False	60
C	True	True	False	True	True	True	False	True	True	True	80
D	True	False	True	False	False	True	False	True	True	True	60
E	True	True	True	True	True	False	True	True	False	True	80
F	True	True	True	True	True	True	True	False	True	False	80
G	True	False	False	False	True	False	True	False	False	True	40
H	True	True	True	False	False	False	True	True	False	True	60
I	True	False	True	False	True	True	False	False	False	False	40
J	True	False	True	False	False	False	True	False	False	True	40
Space	True	False	False	True	True	True	True	False	True	True	70

Table 2 Part 2 of Analysis Table

Alpha-bets	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10	Success Rate (%)
NOTHING	True	True	True	True	True	True	True	True	True	True	100
O	True	True	True	True	True	True	True	True	True	True	100
P	True	True	True	True	True	True	True	False	True	False	80
Q	True	False	False	False	True	False	True	False	False	True	40
R	True	True	True	False	False	False	True	True	False	True	60
S	True	False	True	False	True	True	False	False	False	False	40
Space	True	False	False	True	True	True	True	False	True	True	70
T	True	False	True	False	False	False	True	False	False	True	40
U	True	False	False	True	True	True	True	False	True	True	70
V	True	True	True	True	True	True	False	True	True	True	90
W	True	True	True	True	True	True	True	True	True	True	100
X	True	True	True	False	False	False	True	True	False	False	50
Y	True	True	False	True	True	True	True	True	True	True	90
Z	True	False	False	False	True	False	True	True	True	True	60

VI. CONCLUSIONS

Hand Gesture Recognition is an exceptionally wide and difficult problem, in this project, we have attempted to execute a system that perceives pre-characterized hand gesture with some significant suspicions like simple background, a camera concentrated close by the hand region and decreased the level of movement, which isn't the situation, in actuality. Recognition, then again, requires client interface to record the gesture and after that request that the system classifies it. Aside from this, the system is strong to changing non-extraordinary light conditions, diverse skin hues, distinctive yet straightforward foundation, various paces of movement and distinctive direction of hand while performing gestures.

We explored the different methodologies and the strategies utilized for hand gesture recognition. We found the dataset and the strategies to be utilized as referenced in the Experimentation area as per our suppositions, we picked two procedures for each stage with the exception of the recognition phase and dependent on the outcome we got, we picked one out of them or a combination of both. To actualize these and different procedures, we found out about OpenCV and the various function it gives. Regarding Machine Learning, in spite of the fact that we utilized Lloyd's Algorithm and Support Vector Machine, and an API to actualize the TensorFlow utilizing python.

REFERENCES

1. Elmezain M., Al-Hamadi A., Appenrodt J., Michaelis B., "A Hidden Markov Modelbased continuous gesture recognition system for hand motion trajectory," Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, vol., no., pp.1-4, 8-11 Dec. 2008.
2. Gesture - Wikipedia, the free encyclopedia. 2015. Gesture - Wikipedia, the free encyclopedia. [ONLINE] Available at: http://en.wikipedia.org/wiki/Gesture. [Accessed 19 April 2015].
3. Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research ,Volume 11, pp 290-296 , 2016.
4. Neelam K. Gilorkar, Manisha M. Ingle, "Real Time Detection and Recognition Of Indian And American Sign Language Using Sift", International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 5, Issue 5, pp. 11-18 , May 2014
5. Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015
6. Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people", International Journal of Science and Research (IJSR)