

Hand Gesture Recognition Algorithm to Identify American Sign Language

Rakesh Kumar, Shubham Sinha, Shobhit Dixit
GLA University, Mathura

Article Info

Volume 83

Page Number: 10310 – 10319

Publication Issue:

May - June 2020

Abstract:

Sign language is the most expressive and natural way which is used by hearing and visually impaired people for communication. These people use their hands for communicating with each other, which is not understandable to the normal person. And overall leads the isolation of these peoples from rest of the world. Also, different country has different languages for communication between normal as well as people who have severe hearing impairments. Therefore, the objective of this paper is filling the gap between normal and deaf people through design an algorithm to recognize the different characters of specifically American Sign Language by gesture recognition of the human hand. This proposed algorithm used the information based on the contours and convexity shortcomings to identify the American Sign Language alphabets and digits. The key aim of the proposal is to identify the 17 characters and 6 digits of the American Sign Language without support of any specialized hardware, or any other complex machine learning algorithms. Simulation results validate that, the proposal recognize the American Sign Language alphabets and digits with accuracy more than 85 percent with simple and less complex algorithm within a specified time limits.

Article History

Article Received: 19 November 2019

Revised: 27 January 2020

Accepted: 24 February 2020

Publication: 18 May 2020

Keywords: Sign Language, American Sign Language, Gesture Recognition, OpenCv, Contours, Convexity.

1. Introduction: Sign language is the basic way of communication for the people who are vocally and visually impaired. Normal people never try to learn these languages so this leads to isolation of the deaf people. So our aim is to develop a system that would ease the difficulty of communication between these two groups of the society.

Indian sign language (ISL) uses both hands to represent each alphabet and gesture. ISL alphabets are derived from British Sign Language (BSL) and French Sign Language (FSL). Whereas American Sign Language (ASL) uses single hand for each 26 alphabets which make it easier for computation as compared to Indian sign language (ISL). This is the most growing field of research to translate various sign language as it is reducing the communication gap between the hearing impaired people and normal people. Communication with the hearing

impaired people is a great challenge for the normal people in the society, this can be understanding as there means of communication that is sign language requires an interpreter. Conversion of their gestures to speech as well as text would be a great benefit for these kind of people.

According to World Health Organization till the year 2015, more than 5% of the world population has hearing disability. It is also there that more than 75% of deaf population is unemployed. One of the major reason for this is the ordinary people are feel difficulty in communicating with them. Sign language is one of the basic nonverbal communication solution. It can be done using different body parts, hand shapes, positions and movements of hand, arms, facial expression or movements of lips. A sign language is a language

uses sign or actions to communicate instead of sounds.

Some researchers proposed algorithms to recognize the hand gestures of deaf people. Like in [10], author introduces a special purpose image processing algorithm that have been developed to identify signs accurately from the ASL. One important benefit of this finding is that the person making the signs is not required to use any artificial gloves or markers. This concept is good because it reduces the extra cost of gloves. Many gesture based systems have the common element of markers on the hand [11,12], data gloves or coloured gloves [13] worn by the user to allow the gesture and pose to derived. But all these techniques require specialized hardware and machine learning based image processing algorithms, which makes the solution costlier for common man.

In order to help common people to easily communicate and understand these physically impaired people, this proposed approach translates sign language into text and then into speech. Unlike face detection, a hand detection in live videos against different background is more difficult as the hands change their pattern and orientation rapidly during computation. As compared to faces, the human hands do not have different features and contains many part and joints.

The proposed approach focuses on American Sign Language (ASL), which contains 24 hand signs which are static and 2 gestures which are dynamic, which help to make the whole word letter by letter. Except the letter 'J' and 'Z' all are static in ASL. Based on these 26 alphabets, our proposal is able to identify characters through hand gesture signs shown in figure 1.

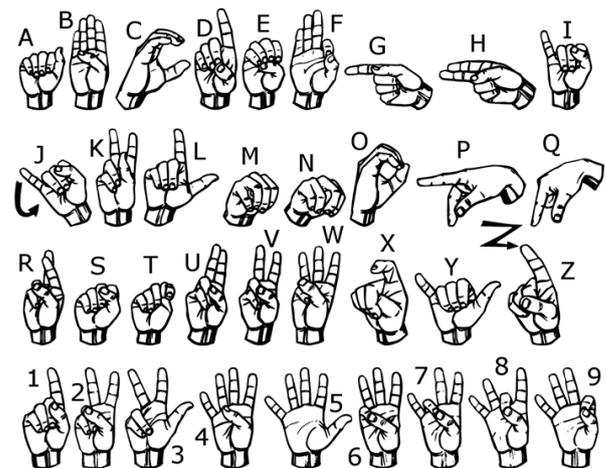


Figure 1: 26 Alphabets and 9 digits of American Sign Language

This paper discusses the recognition of letters from the ASL. Recognition is done using Contour Analysis and Feature Extraction techniques. The goal is to extract human hand properties and various features for the detection of sign language using OpenCv libraries to extract the different features for computing the ASL.

This system will surely contribute towards the society by helping hearing impaired people. The system will work as an important tool for deaf and dumb people to communicate with normal person. The main objective is to build a bridge between normal people and hearing impaired people.

Rest of the paper organize as; section 2 discuss the review of literature proposed by some researchers, section 3 describe the methodology of proposal, followed by performance evaluation in section 3, and lastly section 4 depicted the conclusion and future aspects of the proposal

2. Literature Review: There are varies approaches used by different researchers for the recognition of hand gestures, like glove based approach, vision based approach, artificial neural network, convolutional neural network, and many more. The whole process can be divided into three task: Hand Gesture Analysis, Feature Extraction, and Gesture Recognition.

There are two approaches to collect the hand gesture data one way is vision based approach and another is glove based approach. The vision based approach is further classified into two groups called 3D hand model based approach and 2D hand model approach. A classification approach for sign language recognition was proposed in [1], in which, authors used Zernike moments to identify the orientation of the hand they also considered centre of gravity and calculated it using different formulas. The authors, also used HSV Segmentation and Finger-tip detection, Static Gesture recognition was carried out on a lexicon of 24 alphabets (a-y, excluding j) and it succeeded with approximately 93% accuracy. Another paper[2] used three algorithms for this purpose naming Har_Like_Features, The AdaBoost based learning algorithm, to recognize the hand gesture. A simple OpenCvbased algorithm proposed in [3] to detect the gestures and then find the contours, hull and convexity defect using which 13 characters of ASL.

Similarly, in [4] feature extraction techniques are proposed to identify ASL digits ranging from [0-9] and obtained 82.92% using COHST and wavelet features methods. American sign language(ASL) using back propagation neural network is proposed in[5] and achieved an accuracy ranging from 75-80%. Through the use of raw feature classifier and histogram feature classifier, authors in [6] shows the sign language recognizing of 26 hand gestures in Indian sign language using MATLAB.

The classifier Convolutional Neural Network (CNN) is used to create a classification model [7] that classifies the frames into 26 separate classes representing 26 English alphabets. Lastly, the classification model evaluation is performed with test data providing the output in text or voice type. In [8], many combinations of static gestures are combined with 24 alphabetic symbols (excluding 2 motion gestures J and Z). Through gesture's features of Histogram of Oriented Gradients (HOG) and Local Binary Pattern (LBP) will be extracted from the training images. In [9], newly developed system

for the identification of standardized American sign language comprising 26 English alphabets from 'A' to 'Z' has been introduced. Another glove based contribution proposed in [14], in which, a device is developed with sensor gloves for recording and translating American Sign Language (ASL). The proposal discussed in [15], aims to develop an American Sign Language Detection and Translation to Text based on automatic vision. HSV colour model is used for the detection of the human skin colour from the image. Then applies edge detection to detect the shape of the hand from the image.

The contributions made by researcher discussed above is recognize ASL with support of specialized hardware or some complex algorithms, which makes the approach unusable to common person having less resources. Therefore, in this paper, we proposed a less complex algorithm to fill the gap. Our proposal used defect calculation and angle detection between the fingers and did some calculation to find out the gesture it match with original ASL. The best part of this method is that, it requires little hardware support in comparison to other existing models and make the approach more realistic for daily used personal computers.

In next section, we discuss the proposed framework of the Gesture recognition algorithm in detail, and in last we, analyse the findings and calculate the accuracy of the proposed model.

3. Proposed Hand Gesture Recognition Algorithm for American Sign Language (ASL):

Sign Language is the most natural and expressive way for the deaf and dumb. Normal people never attempt to learn the sign language to interact with the deaf and dumb. This leads to deaf and dumb people becoming segregated. But if the device is programmable in such a way as to convert sign language into text or speech format, that will bridge the gap between common people and the deaf and dumb.

In this work we focus on ASL, because of single handed operations, which overall reduce the

complexity of the algorithm. The proposed algorithm is visual based approach. The image of hand is captured by using openCv and then it is converted into digital representation. Digital image is matrix of scalar or vector values depending upon monochromatic or coloured. OpenCv libraries are used for processing the image and finding the contours. Contours are defined as set of point which describes the boundary of the object. For example, contours of book are rectangle. OpenCv Libraries provide efficient implementation of algorithm to find contours. It provides additional feature of contour extraction in hierarchy, as well as approximation of identified contours. It approximates the line contour with collection of points which is a very useful function and significantly accelerates further contour processing. Such a contour, represented as a series of points that can be enclosed in an n-dimensional polygon also known as hull.

Hull can be concave or convex polygon, as the geometrical shape. We can say the hull is convex when a line inside the polygon cannot be drawn that would cross its border. If so, then the polygon is not convex and thus includes convexity defects. These area descriptive properties, will significantly help with design of algorithm since human hand does have huge convexity defects between fingers.

The framework of the proposed Hand Gesture Recognition Algorithm is shown in figure 2. This framework takes input in the form of video, then convert it to frames for processing and finding digital representation of frames to recognise the hand gesture. After getting the output text and speech representation used for the recognised sign. Various internal steps of the algorithm are described below:

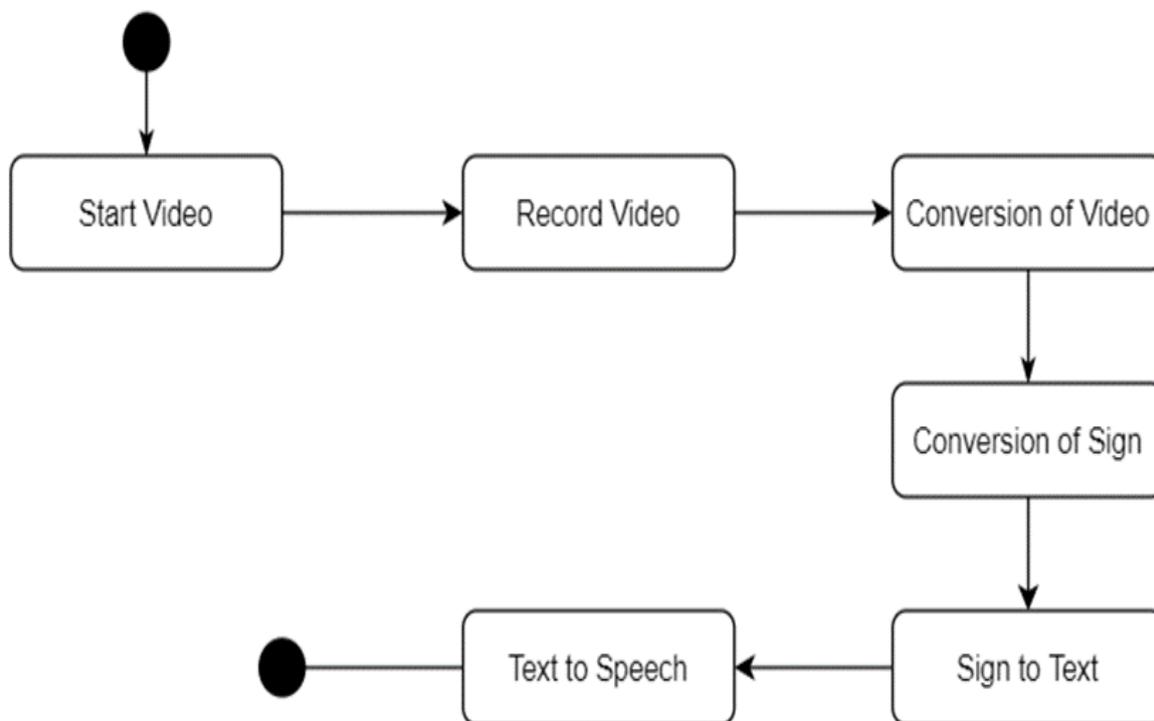


Figure 2: Framework for Hand Gesture Recognition

3.1. Threshold Creation: It is very crucial to create threshold image in hand detection one need to find the skin colour and isolate the foreground from the

background as hand needed to be the region of interest(ROI).

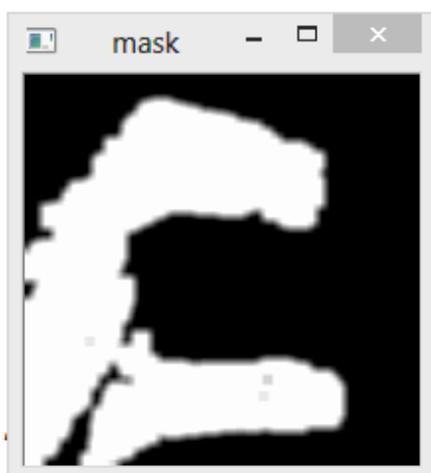


Figure 3: ThresholdImage of Letter C

3.2. Finding Contours: Contours can be defined as the curve joining all the continuous points with same intensity or same colour. Hand is identified by the inbuilt that find contours with the help of OpenCv. The function then returns the array of co-ordinates of the formation of contours.

`cv2.findContours()` function takes three arguments, first one is source image, second is contour retrieval mode, third is contour approximation method. And it outputs the image, contours and hierarchy.

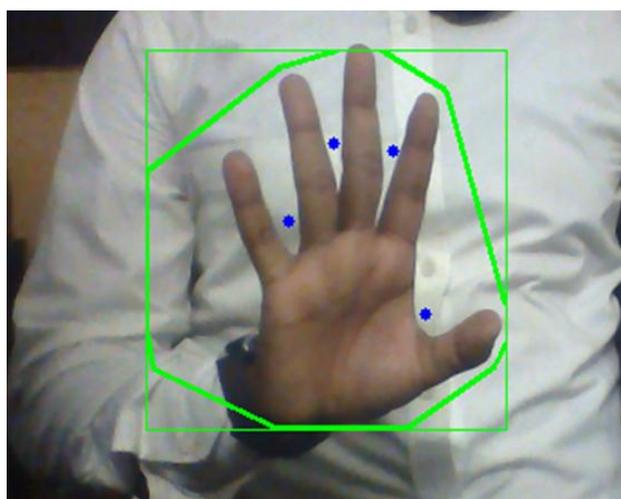


Figure 4: Contour of The Hand with Defects

3.3. Convexity Defect: The data obtained from convexity analysis is used to find the number of convexity defect which are also known as irregularity in the contours. This is the valuable

information because it is used to find number of fingers which is used to find the character it represents.

Number of contours are calculated by the following method for that angle between two fingers, the angle between wrist and the fingers is calculated, the round shape it is made when one closes the fingers are of each polygon which is formed by the hand and shape between the fingers.

3.4. Identification of Letters:

Letter A: For computing and identifying letter A of ASL, the difference between the area of circle and area of contours. The circle is formed by bounding the contours. Following this method, it is found that difference between area of contours and area of circle is very less so this method is used because this difference between areas make the letter A stand out of all the alphabets which is detected.

Letter B: For computing Letter B entire area of contours is calculated as it has largest area among all letters.

Letter C,L,O,V,Y: These letters are identified when condition of letter A fails and the number of defect is one. The angles is calculated between the fingers. These angles are calculated by the OpenCv build in feature which calculate the figures overall orientation the angles and other features based on these angle information letters C,L,O,V,Y are computed.

Letter D,G,H,I,J,U: The data obtained from the OpenCv build in features combination of these information is used to identify these letters. The data obtained which are used are Solidity, angles and the aspect ratio by testing these parameters it is found that these parameters are efficient to find these characters from the ASL.

Letter F and W: These are the only alphabets in the entire American sign language with two convexity

defects. After finding two convexity defects angles are calculated and these two letters are identified.

Letter P,Q: These letters are identified when condition of letter A fails and the number of defect is one. The angles is calculated between the fingers. These angles are calculated by the OpenCv build in feature which calculate the figures overall orientation the angles and other features based on these angle information letters P,Q are computed for these two angle should be greater than 180 and less than 270 degrees.

Digit 0:It is identified by calculating the no of defects if the number of defects are equal to zero and the angle is less than 90 degrees and the contour ratio is also taken in consideration to differentiate between zero and one, for zero the contour ratio is very high then the given gesture is identified as zero.

$$\text{Contour Ratio} = \left(\frac{\text{Area of Hull} - \text{Area of Hand}}{\text{Area of Hand}} \right) \times 100 \quad (1)$$

Digit 1:It is identified by calculating the no of defects if the number of defects are equal to zero and the angle is less than 90 degrees and it has its own contour ratio which is less than that of zero, then the given gesture is identified as one.

Digit 2:It is identified by calculating the no of defects if the number of defects are equal to one and the angle is less than 90 degrees then the given gesture is identified as two.

Digit 3:It is identified by calculating the no of defects if the number of defects are equal to two and the angle is less than 90 degrees then the given gesture is identified as three.

Digit 4:It is identified by calculating the no of defects if the number of defects are equal to three and the angle is less than 90 degrees then the given gesture is identified as four.

Digit 5: It is identified by calculating the no of defects if the number of defects are equal to four and the angle is less than 90 degrees then the given gesture is identified as five. Figure 5(i) to 5(v) represent the threshold of Hand Sign.

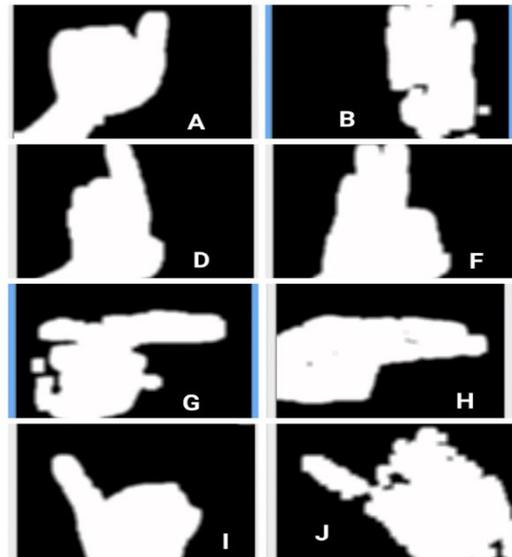


Figure 5(i):Threshold Images for Sign A, B, D, F, G, H, I, and J

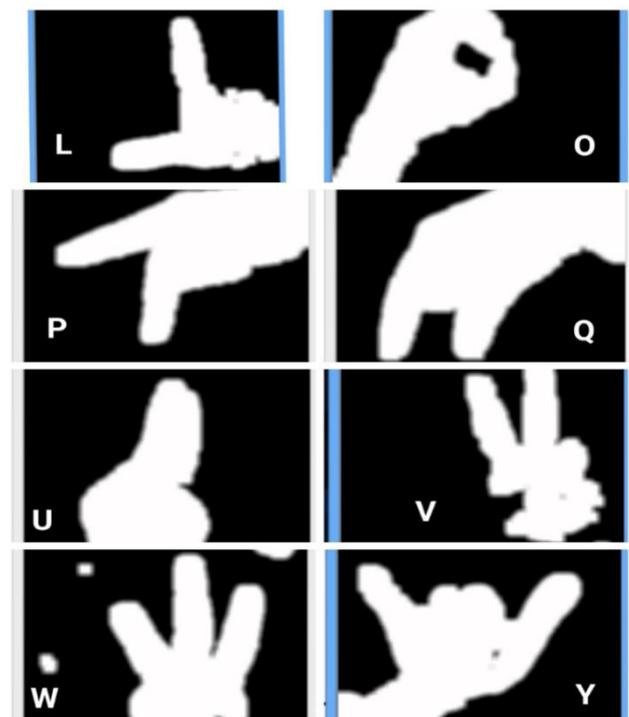


Figure 5(ii):Threshold Images for Sign L, O, P, Q, U, V, W, and Y



Figure 5(iii):Threshold Images for Sign C

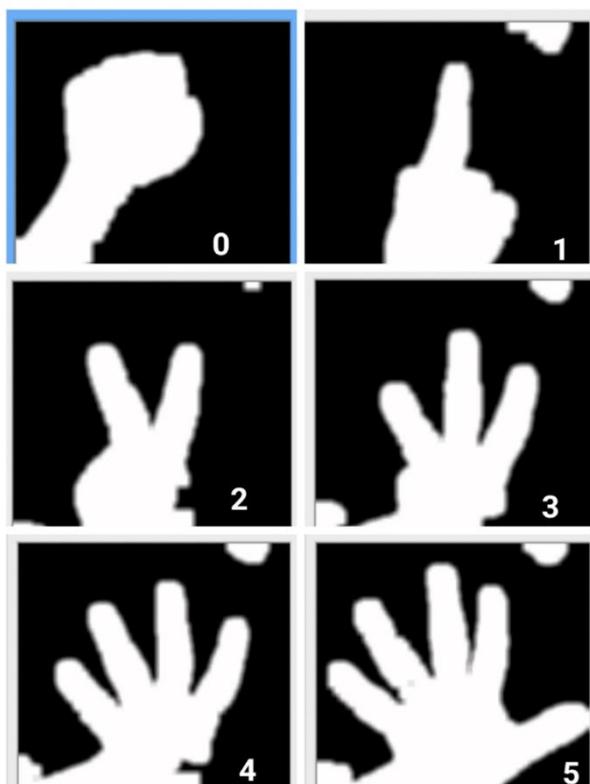


Figure 5(iv):Threshold Images for Digits1 to 5

After recognize sign using OpenCv, output is coming in form of text. By using OpenCv libraries, this text output is converted to speech and provided to the desired user. The proposal uses very less complex procedure to obtained the output in form of text as well as speech. The overall cost in terms of time and space as well as hardware requirements is very less expensive in comparison to existing techniques. In next section we analyse the performance of the algorithm in terms of complexity and accuracy, and validate the findings.

4. Performance Analysis:In this section, the performance of the proposal analyse in terms of accuracy, which is obtained by the ratio by which our algorithm detects the hand gesture. First of all, we find the ratio for each 17 letters and six digits we consider for ASL. Then after find the average accuracy. Results shows that, our approach getting accuracy more than 85%.

In first step, it is important to find our region of interest then calculate the contours and the angles between the palm and the fingers on the basis of which proposed approach has found the 17 ASL characters and then the numbers of defects is calculated in order to find the six digits i.e. (0-5). Figure 6(a) and 6(b) shows the plot for two letters B and V. Area of hand is calculated a blue box is made to make a frame in which hand needed to be placed and the area of the hand is calculated which is shown by the green colour and the defects between the fingers is shown by the red colour. By using all these data this approach gives the satisfactory result which can identify the 17 ASL characters and the six digits.

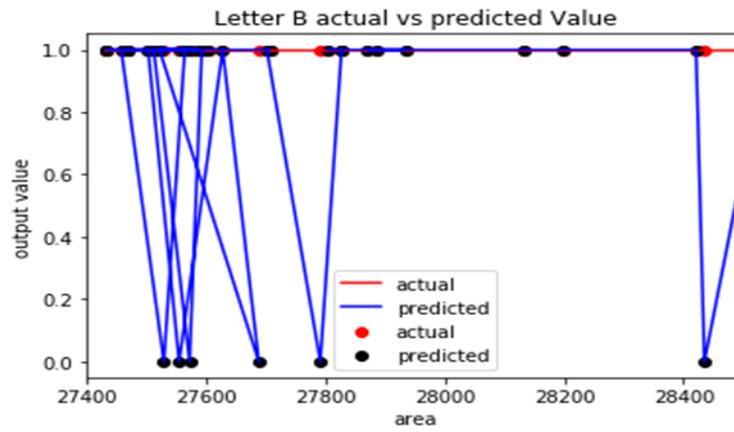


Figure 6(a): Ratio between Actual and Predicted values of Letter B

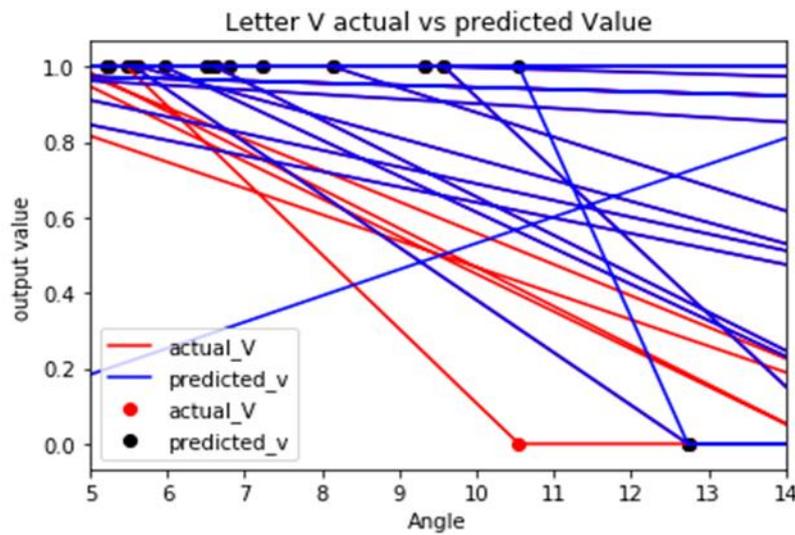


Figure 6(b): Ratio between Actual and Predicted values of Letter V

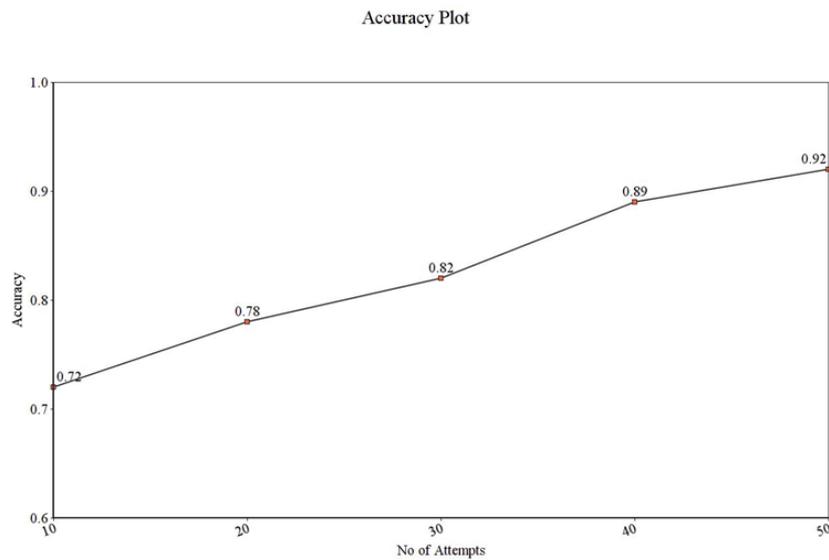


Figure 7: Plot of Average Hand Gesture Recognition Accuracy for various Attempts

In above figure 7, average accuracy of Hand gesture recognition is plotted for each 17 letters and 6 digits for various attempts. It is clearly depicted from figure, for higher attempts, the accuracy is above 90%, which is lower in less attempts. The overall accuracy is more than 85%. Though it runs with an accuracy of 85.41% but still it is efficient because it requires very less hardware support simple hardware requirement which is present in almost all personal use computers.

5. Conclusion and Future Scope: This research carries us fairly closer in the direction of building a sign language recognition system under expected background. The method proposed by us is extraction of Contours and Hull which is defined as finding the boundaries of an object which can be convex or concave. It comprises of successfully fetch informative part from the versatile video background. The system recognizes ASL with complex background from live videos and translate them into text and speech. The typical recognition rate of the overall sign language recognition is 85.41% which is a superior score for the proposed sign language recognition system. Other methods use complex machine learning algorithms and these algorithm requires heavy hardware support to run but our system require minimum hardware requirements that are simple 2gb RAM system with a 2gb graphic card which is more than enough.

There is a vast scope in future for research and implementation in this field. The upcoming years could witness a combinatorial explosion of different methodologies. It can also be developed to work on smart phones so that it can be reached to the greater number of people and easy to use.

References:

1. AnupKumar, KarunThankachan, Mevin M. Dominic, "Sign language recognition", Recent Advances in Information Technology (RAIT), International Conference, 2016.
2. ShivashankaraSs, Dr.Srinath S, "American Sign Language Recognition System: An Optimal Approach", International Journal of Image, Graphics and Signal Processing 10(8), 2018.

3. https://github.com/AnupamYedida/Sign_Language_Recognition/blob/master/ASL%20Recognition.docx
4. Srinath S, Ganesh Krishna Sharma, "Classification approach for Sign Language Recognition", International Conference on Signal, Image Processing, Communication & Automation, 2017.
5. Tülay Karayllan, Özkan Kılıç, "Sign Language Recognition", 2 nd International Conference on Computer Science and Engineering, (2017), pp.1122-1126.
6. Mahesh Kumar N B1, "Conversion of Sign Language into Text", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9 (2018) pp. 7154-7161
7. Prateek S.G., Jagadeesh J., Siddarth R. ; Smitha Y.,P. G. Sunitha Hiremath, Neha Tarannum Pendari, "Dynamic Tool for American Sign Language Finger Spelling Interpreter", 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN).
8. Huy B.D Nguyen ; Hung Ngoc Do, "Deep Learning for American Sign Language Finger spelling Recognition System", 2019 26th International Conference on Telecommunications (ICT).
9. Sriparna Saha ; Rimita Lahiri ; Amit Konar ; Atulya K. Nagar, "A novel approach to American sign language recognition using MAdaline neural network", 2016 IEEE Symposium Series on Computational Intelligence (SSCI).
10. Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari, "Finger Detection for Sign Language Recognition", Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.
11. V. Buchmann, S. Violich, M. Billingham, A. Cockburn "FingerARTips: Gesture based direct manipulation in Augmented Reality", In Proc. of the 2nd international conf. on Computer graphics and interactive techniques in Australasia and South East Asia (Graphite 2004). 15-18th June

- Singapore, 2004, ACM Press, New York, New York, pp. 212-221.
12. Y. Kojima, Y. Yasumuro, H. Sasaki, I. Kanaya, O. Oshiro, T. Kuroda, Y. Manabe and K. Chihara," Hand Manipulation of Virtual Object in Wearable Augmented Reality", In Proc. 7th International Conf. on Virtual Systems and Multimedia (VSMM'01), pp 463-470, October 2001.
 13. Y. Iwai, K. Watanabe, Y. Yagi, M. Yachida, "Gesture Recognition by Using Colored Gloves", IEEE International Conference on Systems, Man and Cybernetics (SMC'96), Vol. 1, pp. 76-81, Beijing, China, Aug. 1996.
 14. KirateyPatil, GayatriPendharkar, Prof. G. N. Gaikwad," American Sign Language Detection", International Journal of Scientific and Research Publications, Volume 4, Issue 11, November 2014 ISSN 2250-3153.
 15. SharmilaKonwar, B. Sagarika Borah, C. Dr.T.Tuithung," An American Sign Language Detection System using HSV Color Model and Edge Detection", 2014 International Conference on Communication and Signal Processing.