DOI: http://dx.doi.org/10.26483/ijarcs.v11i1.6498

Volume 11, No. 1, January-February 2020

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

RECOGNITION OF ACCESSIBILITY FEATURES USING CNN

Jignesh Tailor
Information Technology
MCT's Rajiv Gandhi Institute of Technology
Mumbai, India

Mithil Shah
Information Technology
MCT's Rajiv Gandhi Institute of Technology
Mumbai, India

Siddhika Tibarewala
Information Technology
MCT's Rajiv Gandhi Institute of Technology
Mumbai, India

Nilesh Rathod
Asst. Professor, Dept. of Information Technology
MCT's Rajiv Gandhi Institute of Technology
Mumbai, India

Abstract: America has a deaf population of an estimated 10 lakhs people. The method of communication amongst the deaf community is sign language. The American Sign Language encompasses static and dynamic signs. This paper describes the method to capture the static signs (Which are the alphabets) and then translate that signs into texts. Image processing techniques are applied on these captured images. Upon the completion of the various image processing techniques, the features are relegated by three different techniques. For training dataset convolutional neural network is used. Finally, the interpreted text output for that sign in the English Language is displayed.

Keywords: Image Processing, Convolutional Neural Network, Sign Language Recognition (SLR), American Sign Language

I. INTRODUCTION

Sign language is widely used by people who are unable to verbalize and auricularly discern or people who can auricularly discern but unable to verbalize. A sign language is composed of various gestures formed by different hand shapes, forms of kineticism and orientations of hands or body, or facial expressions. There are various sign languages across the world, each with its own lexicon. These include American Sign Language (ASL) in Northern America, British Sign Language (BSL) in Great Britain, Japanese Sign Language (JSL) in Japan, South African Sign Language (SASL) in South Africa, Indian Sign Language (ISL) in India, etc.

Gestures are used by the auditorily impaired people to express their thoughts. But the use of these gestures is always circumscribed in the auditorily impaired - dumb community, normal people never endesvor to learn sign language. This causes an immensely colossal gap in communication between the deaf - dumb people and the mundane people. Usually deaf people seek the help of sign language interpreters for translating their thoughts to normal people and vice - versa. But these systems are very costly and does not work throughout the life period of a deaf person. So, a system that automatically recognizes sign language gestures is obligatory.

II. LITERATURE SURVEY

For the American Deaf Community, the usage and learning of Sign Language is restricted to a handful. In order to make translation from sign language to text an authenticity, without any extravagant hardware requirements, the technique of image processing is used.

Fundamentally there are two approaches for sign recognition vision - based and sensor - based gesture recognition [3]. Lots of study has been done on sensor-based approaches like gloves, wires, helmets etc [4][5][6]. But due to the disadvantage of wear it perpetually is not possible, consequently further work is concentrated on Image based approaches [1].

Some previous work has been done on image based approaches for hand gesture and sign recognition in the last few years[14].There has been various methodologies for gesture recognition like HMM(Hidden markov model)[15], ANN(Artificial neural network)[16], Eigen value based[17], perceptual color based[1][18].In [19] authors proposed gesture apperception algorithm using GMM and HMM. Techniques like SVM (Support vector machine) proposed for relegation and particle filtering [20].

Indian Sign Language (ISL) consists of word level signs as well as fingerspelling. Fingerspelling is utilized for letter by letter signing. It is used to code a word for which the sign does not subsist or to accentuate a particular conception or thought. This paper proposes a method for automatically recognizing the fingerspelling in Indian Sign Language. Sign language recognition methods are mainly relegated into two broad categories: contrivance-predicted methods and vision-predicted methods. contrivance-predicted approach needs special hardware devices to extract the physical features of the hand sign such as dimension, angle, kineticism and colour. In comparison, vision-predicted methods use image processing algorithms to detect, track and interpret hand signs. This approach has the advantage that the user does not have to wear maladroit contrivances.

Color Models:

RGB Values

R > 95 and G > 40 and B > 20 and Max {R, G, B} > 15 and mod (R - G) > 15 and R > G and R > B

YCbCr Values:

Using the following equations:

Y = 0.299R + 0.587G + 0.114B,

Cb = 128 - 0.168R - 0.331G + 0.5B,

Cr = 128 + 0.5R - 0.418G - 0.081B.

The contour based potential energy is calculated as follows: Extract the boundary of segmented binary image. Column vector C, where each element in this vector is the weighted sum of the pixel values in the corresponding column of the contour image. The row vector and the column vector are the potential energy to the bottom and left border and is known as the two-dimensional potential energy [2].

III. PROPOSED METHODOLOGY

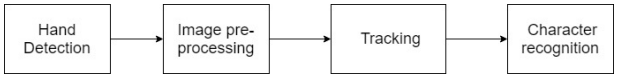

FIG.1 THE OVERVIEW OF PROPOSED SYSTEM

The proposed system comprises of 4 stages. They are as per the following: Hand Detection, Image Pre-processing, Tracking and Character Recognition. The training dataset is performed using Convolutional Neural Network. The TensorFlow and OpenCV libraries are utilized for GUI and to display the recognized American sign language character.

A. Acquisition of Data (Camera Interfacing)

This is the most essential step in sign recognition process. Camera interfacing is consequential errands so as to capture an image with the assistance of a webcam. Gestures can be captured using in built camera to detect hand forms of kineticism and position. Capturing 30 fps will be adequate to process images. Higher quality image may prompt higher computational time and will make system exceptionally moderate and helpless.

B. Image Pre-processing

Image pre-processing contains abstraction of unwanted noise, brightness adjustment and contrast of the image, cropping the image as per the requisite [7]. These are the steps which makes captured image more efficient.

1) Image Enhancement

Images are captured by webcam which is RGB images, yet the solemn issue with RGB images are that they are delicate for different light conditions. Therefore, the RGB data is converted into YCbCr. Where Y is the luma component which signifies luminance data of image and Cb, Cr are chromo components which give shading data of image red distinction and blue difference. Luminance component may make issues so just chrominance components get processed further. After this enhanced YCbCr image is converted into binary image.

2) Color Filtering and Skin Segmentation

In real time, the image captured by webcam contains amassment of frames. There is need to convert RGB images into HSV images, because it is cognate to human color perception. Fundamentally the color spaces differentiate into three components: hue(H), saturation(S), value(V). Image segmentation is typically performed to locate the hand object and boundaries of images, for this HSV color model features avails user to designate boundary of skin color in terms of hue, saturation and value. Value gives effulgence information so therefore it is facile to relegate skin color and non - skin color information in images [8]. In this approach, value of HSV is within range of 0 to 255 to extract and get precise boundary of object.

3) Noise Removal: Gaussian Filtering Method

Gaussian filtering is used to blur images and remove noise and make it more detailed. In one dimension, the gaussian function is:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

FIG.2 ONE-DIMENSIONAL GAUSSIAN FILTERING FORMULA

In the above function σ is standard deviation of the distribution. The distribution is postulated to have a mean of 0. The standard deviation of the Gaussian function plays a consequential role in its demeanor. The values located between +/- σ account for 68% of the set, while two standard deviations from the mean account for 95% and three standard deviations account for 99.7%.

While working with the processed images we have to use two-dimensional Gaussian function. This is simply the product of two 1Dimensional Gaussian functions and is given by:

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

FIG.3 TWO-DIMENSIONAL GAUSSIAN FILTERING FORMULA

A graphical representation of the 2Dimensional Gaussian distribution with mean (0,0) and σ = 1 is shown in the diagram below:

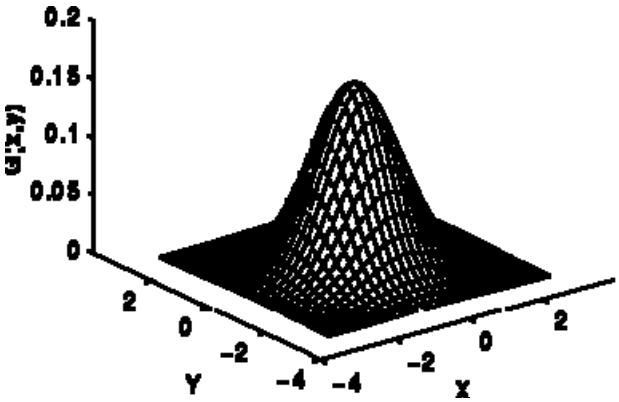

FIG.4 TWO-DIMENSIONAL GAUSSIAN DISTRIBUTION FUNCTION

It works by utilizing 2D distribution as a point - spread function. This is achieved by convolving the 2D

Gaussian distribution function with the pre-processed image. The Gaussian filter is a non - uniform low pass filter. The kernel coefficients diminish with increasing distance from the kernel's centre. Central pixels have a higher weighting than those on the periphery. Larger values of σ engender a wider peek which rudimentary denotes more preponderant of the pre-processed image. Kernel size must increase with increasing σ to maintain the Gaussian nature of the filter. Gaussian kernel coefficients depend on the value of σ. At the edge of the mask, coefficients must be proximate to 0. Gaussian kernel is separable, which sanctions expeditious computation [9]

4) ***Image Segmentation: Thresholding Method***

It is the simplest method of image segmentation. From a grayscale image, thresholding process can be habituated to engender binary images. In thresholding each pixel in image supersede into black pixel, if image intensity is more preponderant than constant value. A primary property which pixels in image can-apportion is its intensity. Hence in thresholding images separate into regions depending on Light and dark regions.

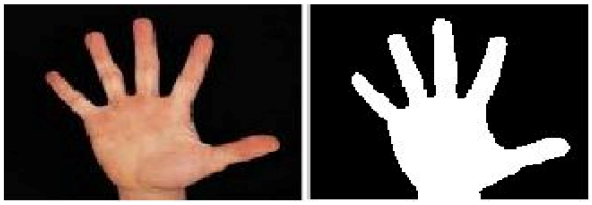

FIG.5 A) INPUT IMAGE B) SEGMENTED IMAGE

5) ***Image Analysis: Blob Detection***

In the field of computer vision, blob detection method is referring to detection of regions in the image which either effulgent or more tenebrous than the circumventing region, fundamentally this technique is characterized as an accumulation of pixels composed into a structure. It is apperception of points in the processed image which varies in features like brightness and color.

6) ***Contour Detection***

In contour detection convexity hull algorithm uses for drawing contour around the palm and finger points detection. In convexity hull algorithm adaptive boosting algorithm use for hand detection. Initial step of convexity hull algorithm is to segment image in which hand is located. For this some feature must be surmised. Here surmised shape of hand but that may change according to the kineticism of hand. consequently, skin color of hand is considered, because it is a way to scale and movement of hand. The next phase of a tracking system contains disuniting hand pixels from non - hand pixels. After segmentation contour is extracted [10].

FIG.6 CONTOUR EXTRACTION [11]

In convexity hull algorithm initially computes maximum and minimum x and y coordinate point and by joining those points form bounding rectangle which contains hull. Like hulls there are other points withal present i.e. convex defects of hand, which are present in between valley of two fingers [12] as shown in fig.7 then by taking an average of all such defects points surely get a center of palm. So, radius of palm considered a depth of palm only. Then ratio of palm radius and distance of hull point from the center point of palm should be more or less to determine finger opening and closing position.

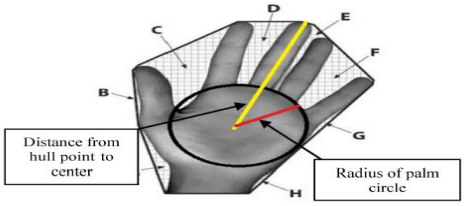

FIG.7 CONVEX HULLS AND DEFECTS [13]

The points which join the boundary of hand are called as convex hulls and letters A, B, C, D, E, F, G, H denotes gaps between fingers are called convex defects. Hence for apperceptions of fingers position, distance(D) of convex hulls from center point should be more preponderant than the radius(R) of inner defects circle. Depending on the value of D and R finger opening and closing can be tenacious. Therefore, Convexity hull algorithm is very convenient and felicitous method for finger point detection.

IV. IMPLEMENTED RESULTS

The dataset collected of the American Sign Language has in it 26 Alphabets and 3 Special Characters which is used to write a complete group of words making sense with the avail of more preponderant, stronger and more complete system. The dataset collected of American Sign Language has 3000 images for every character with all types of possible amalgamation of lights, camera angles. The number of steps required to train entire dataset is 1500 which gives accuracy of 89%. The time required to train the dataset first time is about 2-3 days. After that it only takes 2-3 hours to train the dataset. Bottlenecks of dataset is engendered of every image of dataset. The bottlenecks are then compared to the processed image.

The system is kept live so that the entire sentence is indited with the avail of special characters. The system will not cease the operation until the user is manually close the system. There are 2 screens in which one is for the webcam or any other

input device and the other one is command prompt in which the user can optically discern the output character of American sign language. Both the screens are discrete such that the user can transmute the size of screens manually in terms of width and height.

V. CONCLUSIONS

The American Sign Language is the only communication method between auditorily impaired people of America with the mundane people of America. The Sign Language plays a consequential role in auditorily impaired people life and they have to use sign language for their communication purposes. In this paper, we have built a live system, which can recognize 29 characters of American Sign Language which includes 26 alphabets and 3 special characters. The amalgamation of image processing techniques makes the system more efficient. Using this system, the deaf community can inscribe the entire sentence with the help of American sign language gestures. The accuracy achieved makes it more efficient.

VI. ACKNOWLEDGEMENT

We would like to thank MCT's Rajiv Gandhi Institute of Technology for providing research facilities and to Prof. Nilesh Rathod (Information Technology) department for their support and encouragement.

VII. REFERENCES

[1] V. K. Verma, S. Srivastava, and N. Kumar, "A comprehensive review on automation of Indian sign language," IEEE Int. Conf, Adv. Comput. Eng. Appl.,Mar 2015 pp. 138-142.

[2] Adithya V, Vinod P R, Usha Gopalakrishnan,"Contour potential energy based method for indian sign language recognition", International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013.

[3] A. Chonbodeechalermroong, "Dynamic contour matching for hand gesture recognition from monocular image," IEEE 12th Int. Joint Conf, Comput .. Säe and Soft. Engg. 2015pp.47-51.

[4] Ramirez-Giraldo, D., et al. "Kernel based hand gesture recognition using kinect sensor." Image, Signal Processing, and Artificial Vision (STSIVA), 2012 XVII Symposium of IEEE,2012.

[5] Zhu, Chun, and Weihua Sheng. "Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living." Systems, Man and Cybernetics, Part A: Systems and Humans,IEEE, Transactions on (2011):569-573.

[6] Camastra, Francesco, and Domenico De Felice. "LVQ-based hand gesture recognition using a data glove." Neural Nets and Surroundings.Springer Berlin Heidelberg, 2013. 159-168.

[7] M. Panwar, "Hand gesture based interface for aiding visually impaired," Proc. IEEE Int. Conf. Recent Adv. Comput. Softw.Syst. RACSS 2012, pp. 80-85.

[8] Aliaa A. A.Youssif, AmalElsayedAboutabl, HebaHamdy Ali, "Arabic Sign Language (ArSL) Recognition System Using HMM ", (lJACSA) International Journal of Advanced Computer Science and Applications, Vo1.2, Issue. 11,2011

[9] L. Gu, X. Yuan, and T. Ikenaga, "Hand gesture interface based on improved adaptive hand area detection and contour signature," IEEE Int. Symp. Intel!. Signal Process. Commun. Syst. (ISPACS 2012), no. Ispacs, pp. 463--468

[10] H. Y. Lai and H. J. Lai, "Real-Time dynamic hand gesture recognition,"IEEE Int. Symp. Comput. Consum. Control,2014 no. 1,pp. 658-661.

[11] P. Suresh, N. Vasudevan, and N. Ananthanarayanan, "Computer-aided interpreter for hearing and speech impaired," Proc. IEEE 4th Int. Conf. Comput. Intell. Commun. Syst. Networks, CICSyN 2012, pp. 248-253.

[12] B. Toni, J. Darko, and Ipro, "A robust hand detection and tracking algorithm with application to natural user interface,"2012 pp. 1768-1774.

[13] S. Ganapathyraju and D. Ph, "Defects to Control an IndustrialRobot," 2013 pp. 63-67.

[14] A. K. Tripathy, D. Jadhav, S. A. Barreto, D. Rasquinha, and S.S. Mathew, "Voice for the mute," pp. 2-7, 2015.

[15] Lang, S., B. Marco and R. Raul," Sign language recognition using Kinect" In: L.K. Rutkowski, Marcin and R. T. Scherer, Ryszard Zadeh,Lotji Zurada, Jacek (Eds.), Springer Berlin Heidelberg, pp:394-402,20 11.

[16] S. Saengsri, V. Niennattrakul, and C..A. Ratanamahatana, "TFRS: Thai Finger-Spelling sign language recognition system", Second International Conference on Digital Information and Communication Technology and it's Applications(DICT AP) ,lEEE20 12,pp.457 -462.

[17] Joyeeta Singha and Karen Das, "Indian sign language recognition using Eigen value weighted euclidean distance based classification technique", International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2 , 2013.

[18] Manigandan M. and T. M Jackin, "Wireless Vision based Mobile Robot control using Hand Gesture Recognition through Perceptual Color Space", IEEE International Conference on Advances in Computer Engineering, 2010, pp. 95-99.

[19] Song, Yang, et al. "A Kinect based gesture recognition algorithm using GMM and HMM." Biomedical Engineering and Informatics (BMEJ),20 13 6th International Conference on. IEEE,2013.

[20] Majida Ali Abed , Ahmad Nasser Ismail and Zubadi MatizHazi,"Pattern recognition using Genetic Algorithm",International Journal of Computer and Electrical Engineering, VoI.2,No.3,]une,2010