International Journal of Innovative Research in Computerand Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

Automatic Indian Sign Language Recognition using Sobel Edge Detection

Shrikant Singh [1], Devendra Rewadikar [2], Ankur Taneja [3]

PG Student, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[1]

Assistant Professor, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[2, 3]

ABSTRACT: Sign language is a visual language with its own grammar and gesture that bit differs from the spoken language. It is a hand and facial expression based language that generally used by dumb and deaf people to communicate with each other. Automatic sign language recognition is a challenging task where sign can be recognized by its gesture and body posture. Sign language is different in various countries with its own gesture assigned for visual communication. There are various methods that can achieve goal but the only difference is the precision rate that is directly proportional to the correct recognition and error rate. System can be prefabricated using Sobel Edge Detection with morphological dilation in an effective manner. Sobel Edge Detection is a modern edge detection tool that can sharply extract the outlines of any gesture in an image. Technically, it is a discrete discriminant operator, computing an estimate of the gradients of the image intensity function. At each point of the image, the result of the Sobel – Feldmann operator is either the corresponding gradient vector or the ideal of this vector. System uses various preprocessing or filters to enhance the subject visibility through which a gesture can easily recognizable. System achieved the accuracy as 92.00 % which is bit higher than the previous implementations.

KEYWORDS: Sign Language Recognition, Indian Sign Language, Sobel Edge Detection, Morphological Dilation, Discriminant Operator.

I. INTRODUCTION

Sign language recognition is a breakthrough for helping deaf and dumb people and has been researched for many years. Unfortunately, each research has its own limitations and is still unable to be used commercially. Some research has recognized sign language to be successful, but commercialization requires an expensive cost. Nowadays, researchers have paid more attention to develop sign language recognition that can be used commercially. Researchers conduct their research in various ways.

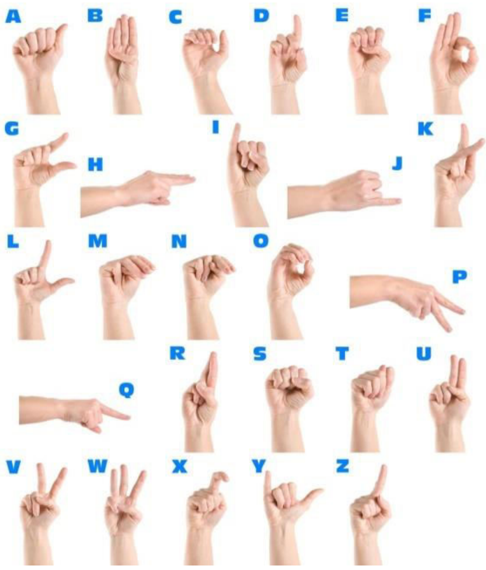


Fig. 1. Indian Sign Language Postures [1]

It starts with data acquisition methods. The data acquisition method varies due to the cost required for a good device, but the commercialization of sign language recognition systems requires inexpensive methods. Methods used to develop sign language recognition also differ among researchers. Each method has its own strengths compared to other methods and researchers are still using different methods in developing their own sign language recognition. Each method also has its own limitations compared to other methods. The purpose of this paper is to review sign language recognition approaches and to find the best method used by researchers. Therefore other researchers can gain more knowledge about the methods used and develop better sign language application systems in the future.

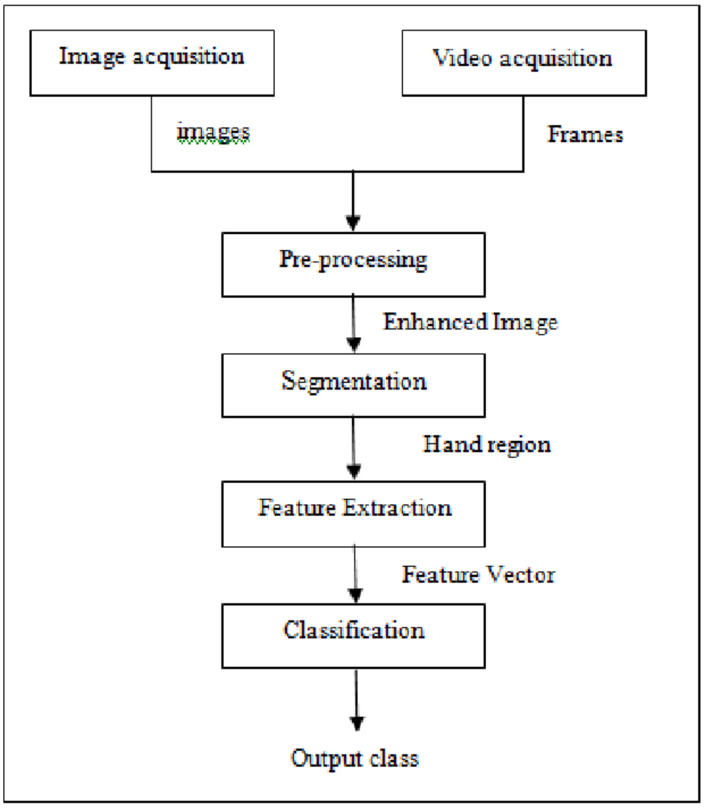


Fig. 2. Automatic Sign Language Recognition [2]

In every traditional approach the first phase is to acquire the frames and pre-process that for image enhancement later segment the hand gesture for illustrating hand gestures that can classifies the sign as per the posture language. Therefore, the need for a computer-based intelligent system is in high demand for the dumb community which will enable them to communicate with all others using their natural hand gestures. This paper presents a method for automatic recognition of signals based on shape-based features. For segmentation of the hand region from images, Otsu's threshold algorithm is used, which chooses an optimal threshold to reduce the square variance within the threshold black and white pixels.

II. RELATED WORK

Bhumika Gupta et al. [3] proposed a system which is based on K-nearest correlation for extracting features. System tried to recognize all alphabets by their gestures using double hand. But Classification method is generally used for classifying regular or irregular objects from an image to better predict the object model. Hand gestures or sign languages are training based data models where supervised learning is important for precise recognition. Segmentation is not a method that can built sign language recognition system because it requires deep analysis what gesture it belongs. G.Anantha Rao et al. [4] proposed a system which is based on CNN. The system depends upon the training samples and machine iterations. The weakness of CNNs lays within the amount of knowledge you provide to them. If you provide them with less, expect the CNNs to perform poorly. CNNs have many parameters and with small dataset, would run into an over-fitting problem because they needs massive amount of knowledge to quench the thirst. So, you give much data, CNNs is more strong and willing to offer you better performance, but when you give less data then CNNs becomes extremely weak. Large dataset may increases the time and space complexity that may requires much time to execute and respond poorly. Snehal Madhukar Daware et al. [5] proposed a system which is based on morphological operations. Morphological operation is an image processing tool that can work with grayscale image or binary image. It filters the image either by eliminating the unnecessary pixels that is called erosion or filling pixels where it loses that is called dilation. But by using this operation or algorithm a complete gesture cannot be portrayed because it can only have active pixels and erode the rest one, it means that it transform or segment an image in black and white pixels (either 0 or 1) that does not dignify the hand gesture.

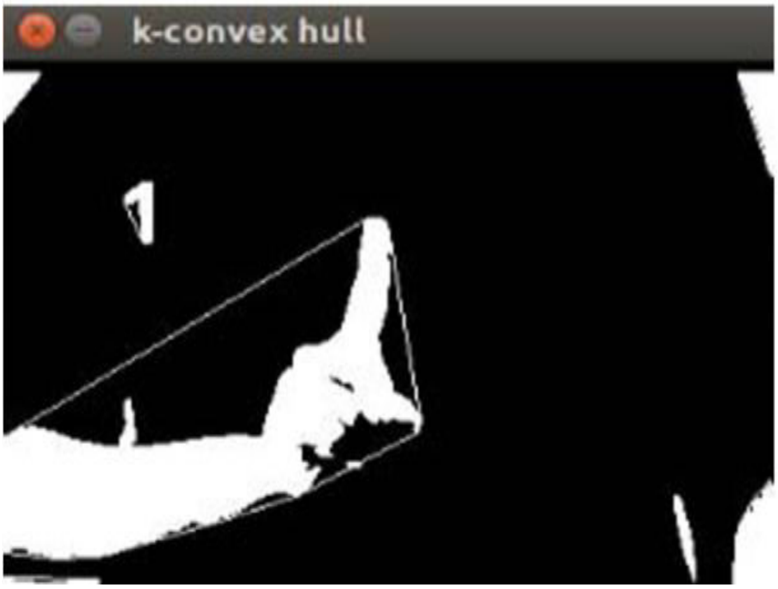


Fig. 3. K-convex hull algorithm [10]

Muthu Mariappan H et al. [6] proposed a system which is based on preprocessing and classification techniques such as morphological operations and fuzzy classification resp. Clustering is known as the process of grouping of similar data items together, while the items in the other clusters are as dissimilar as possible. In fuzzy clustering, the data items may belong to more than one cluster. Among several fuzzy clustering algorithms, fuzzy c-means clustering (FCM) algorithm is used most widely, and this can be used for both supervised learning and unsupervised learning, depending upon the needs. Clustering after erosion may erode the useful information from an image that may distort the sign language gesture which results inaccuracy. K. Revanth et al. [7] proposed a system which is based on Support Vector Machine for classifying sign language gestures from an image. System also uses skin masking technique for segmenting region of interest and eliminating the background by focusing the skin color. For skin segmentation; system uses OpenCV. On perceptual tasks (vision, speech and so on), they are massively outclassed by deep neural networks. On structured data, they are outperformed by gradient boosted trees. Other disadvantages include not giving class probabilities and being rather cumbersome for multiclass problems (you need one model per class). Sruthi C. et al. [8] proposed a system which is based on deep learning methodology i.e. CNN. This work addresses Indian signing static alphabet recognition problem with a vision based approach. A Convolutional Neural Network (CNN) which may be a deep learning technique is employed to make a model named signet, which may recognize signs, supported supervised learning on data the entire process are often divided into CNN training and model testing. A concentric layer consists of a set of matrices, multiplied by the output of the previous layer in a process called convention, in which these features may be core features to detect certain features (such as edges, Color grade or pattern) or complex one (e.g. shape, nose, normal) or mouth). So, those matrices are called filters or kernels. Sign language recognition is a complex issue because it may have similar gestures for different sign languages. A strong learning module is required to achieve the target with high level of accuracy. Salma Hayani et al. [9] proposed a system that uses Convolutional Neural Network (CNN) for recognizing sign language gestures. System recognizes Arabic sign languages of different categories of letters and numbers. Most of the signs may have similar gestures that seriously confuse the system to obtain the correct result. Template matching may return false result when machine has not been trained with sufficient samples and system is not effective for complex gestures and back propagation is not effectively work with CNN that is why the correct gesture recognition rate may degrade due to ineffectiveness.

International Journal of Innovative Research in Computerand Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

III. PROBLEM IDENTIFICATION

Most of the system uses machine learning methods to train the system with various samples for a single hand gesture. But a large dataset can consume the large amount of memory that increases the execution time where it is very important to communicate as earlier as possible with high level of accuracy. Some of the system uses morphological operation which is a weak technique for recognizing any sign language gesture because a gesture may have various distinguish internal edges that reflect those posture uniquely but eroding the image may eliminate or mask the gesture completely that increases the incorrect recognition that indirectly affect the precision rate. In future a system can be developed that may have good accuracy rate with less false alarm or recognition that acquire less execution time. System uses Convolutional Neural Network (CNN) for recognizing sign language gestures. System recognizes Arabic sign languages of different categories of letters and numbers. Most of the signs may have similar gestures that seriously confuse the system to obtain the correct result. Template matching may return false result when machine has not been trained with sufficient samples and system is not effective for complex gestures and back propagation is not effectively work with CNN that is why the correct gesture recognition rate may degrade due to ineffectiveness. CNN do not encode the position and orientation of the object into their predictions. They completely lose all their internal data about the pose and the orientation of the object and they route all the information to the same neurons that may not be able to deal with this kind of information. Accuracy is often important for proper communication. Here the accuracy for correct recognition is 90.03 % which is bit lesser [9].

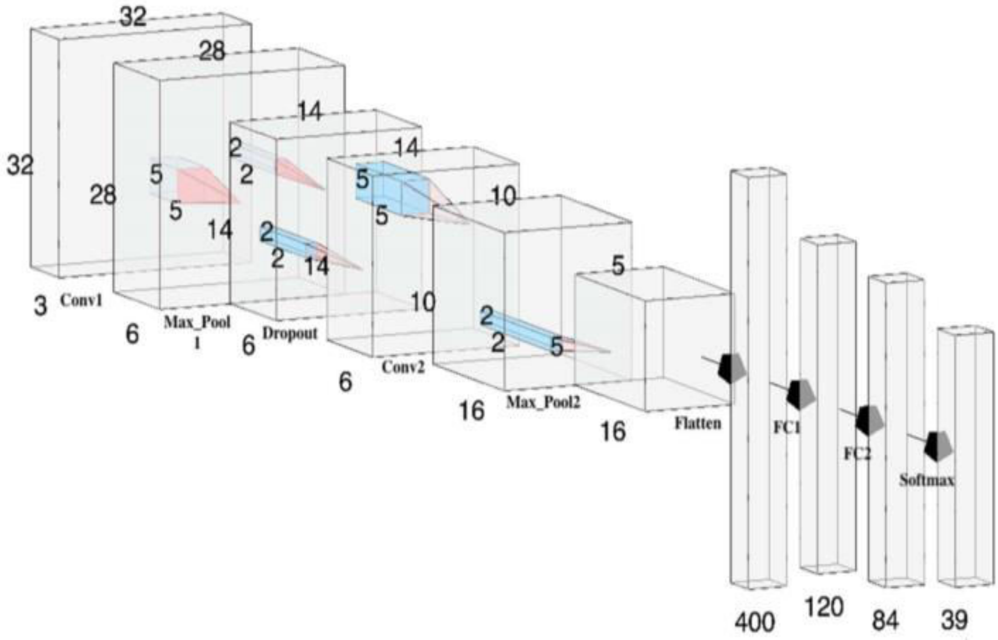


Fig. 4. System Architecture [9]

IV. PROPOSED WORK

Here the proposed work is able to recognize Indian sign language with higher prediction rate and less false acceptance rate. System purely debates gesture based hand sign recognition that is helpful for dumb and deaf people while having conversation. System uses Sobel Edge Detection and morphological dilation that deals with best precision. System also uses histogram equalization for adjusting the contrast of an image for better segmentation of region of interest. A histogram of an image is a graphical representation of intensities. In simple terms, it has been assumed that at each intensity value; it represents the number of pixels. The color histogram graph of an image represents the number of pixels for each color components. The histogram equations cannot be applied separately for red, green, and blue components because it causes dramatic changes in the image's color balance. However, if the image firstly converted to another color mode, such as the HSL / HSV color space, then the algorithm can be applied to the luminance and the value channel as a result of changes over the color and saturation of the image.

A. Sobel Edge Detection

Sobel is a filter in image processing for extracting edges or slight changes in pixel intensities. Sobel is a digital filter which helps to extract edges on the basis of angular matrix. Lesser the value of gray level darker the area and bigger the value of gray level lighter the area. The color range lies between 0-255. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. By the help of Sobel edge detection an object can be detected from an image with high level of accuracy. Sobel is better than all other edge detection technique due to its edge detailing. Consider A as an input matrix -

A =

a11	a12	a13	…	
a21	a22	a23	…	
a31	a32	a33	…	
…	…	…	…	

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A, \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

Where A as an input 2D image array or matrix, G_x and G_y are the gradient kernel that will be multiplied with input image array A. Where G_x is the horizontal gradient and G_y is the vertical gradient. Negative gradients appear darker, and positive gradients appear brighter. Computing the value at each pixel and shifting the row towards right till the end row has been reached. The example below shows the calculation of a value of G_x:

a11	a12	a13	…	
a21	a22	a23	…	
a31	a32	a33	…	
…	…	…	…	

Input Matrix

b11	b12	b13	…	
b21	b22	b23	…	
b31	b32	b33	…	
…	…	…	…	

Output Matrix

Kernel =

+1	0	-1
+2	0	-2
+1	0	-1

$$b_{11} = a_{11}*1+a_{12}*0+a_{13}*(-1)+a_{21}*2+a_{22}*0+ a_{23}*(-2)+ a_{31}*1+ a_{32}*0+ a_{33}*(-1)$$

Similarly, each pixel will be calculated according to the kernel matrix and finally G_x has been computed.

The example below shows the calculation of a value of G_y:

Kernel =

+1	+2	+1
0	0	0
-1	-2	-1

$$b_{11} = a_{11}*1+a_{21}*0+a_{31}*(-1)+a_{12}*2+a_{22}*0+ a_{32}*(-2)+ a_{13}*1+ a_{23}*0+ a_{33}*(-1)$$

At each pixel in the image, the gradient approximations given by Gx and Gy are combined to give the gradient magnitude, using:

$$G = \sqrt{G_x^2 + G_y^2}$$

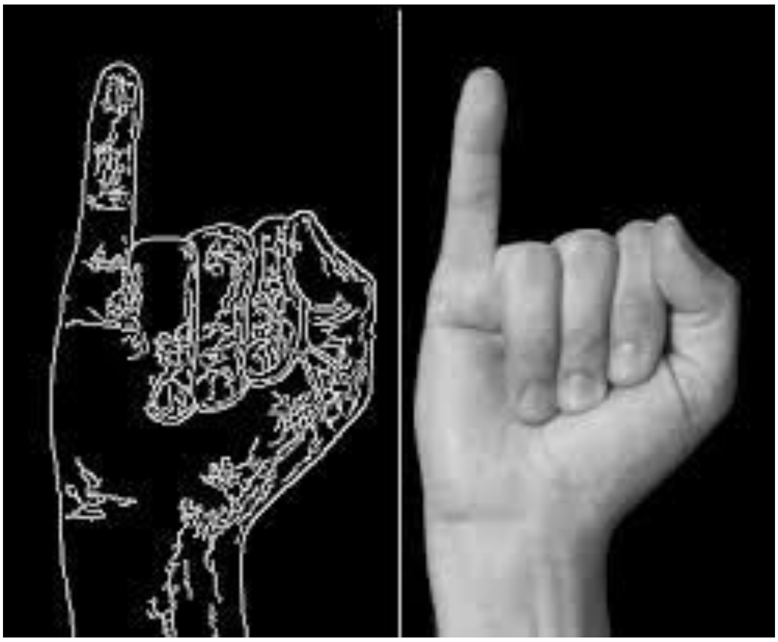


Fig. 5. Sobel Edge Detection of Sign Language Gesture

Morphological operation is an image processing technique that process images on the basis of their shape and size. In morphological operations, each pixel values of output image comes out from pixel values of input image by comparing it with corresponding pixels with neighbors. Generally there are two main operations in morphological function – erosion and dilation. Erosion is an operation where pixels are removed from the boundaries of an image where extra regions may belong whereas dilation adds the pixels to the boundaries where pixel may required. Morphology is a collection of non-linear operations which is related to the shape or features of an image. Morphological operations are depend on the relative order of pixels, not on their numerical values or coordinates, and are therefore particularly suited to the processing of binary images. Morphological operations can also be applied over grayscale images, such that their brightness functions are unknown and therefore have no or slight significance in their absolute pixel values.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

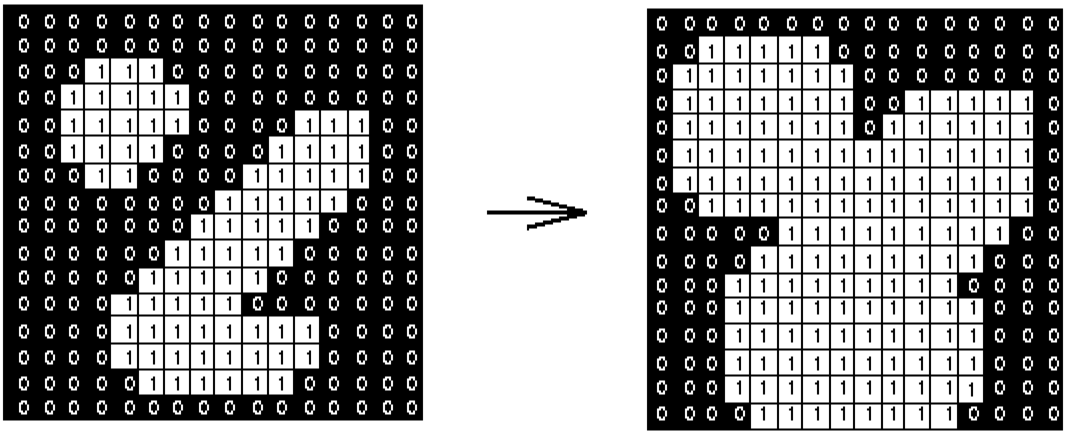


Fig. 6. Morphological Dilation [11]

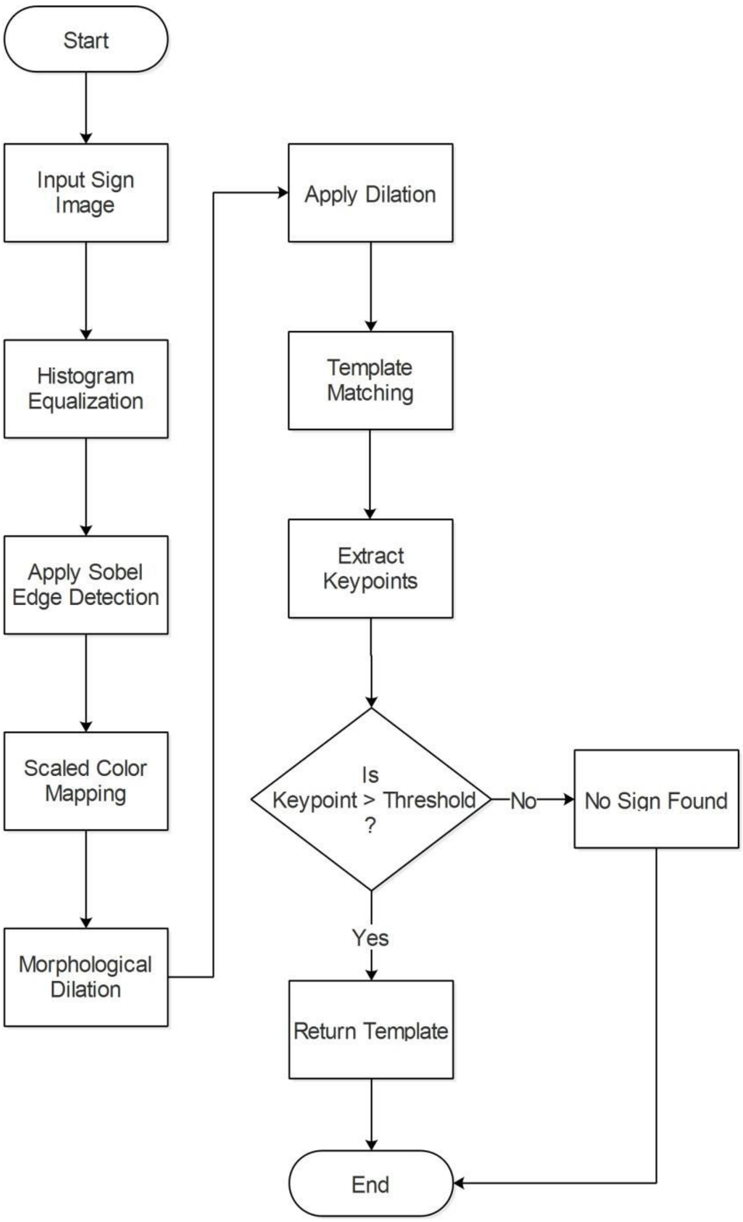


Fig. 7. Flow Chart

A. Sobel Magnitude & Dilation Algorithm:

Require: $G_x \leftarrow$ Horizontal mask, $G_y \leftarrow$ Vertical Mask, x, y $\leftarrow$ coordinates, G $\leftarrow$ Absolute magnitude, A $\leftarrow$ Input image, x $\leftarrow$ Grayscale image, T_r is a threshold and $P_x \leftarrow$ Probability.

INPUT: A $\leftarrow$ Input sign image as 2D array

OUTPUT: H $\leftarrow$ Absolute gradient magnitude

Step 1: Input 2-dimentional image as array

Step 2: Convert RGB image to grayscale

Step 3: Adjust contrast using histogram equalization

$$cdf_x(i) = \sum_j^0 P_x(j)$$

where cdf is cumulative distribution function, x is grayscale image, i is gray levels and P is probability

Step 4: Apply Sobel using gradient mask G_x & G_y, where G_x is horizontal mask & G_y as vertical mask.

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A, \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

$$G = \sqrt{G_x^2 + G_y^2}$$

Then compute, G^{-1} for smoothening extracted matrix for clustering.

Step 5: Apply Dilation by multiplying each pixel with scale factor $\frac{1}{2}$

$$(x, y) \rightarrow \left(\frac{1}{2} \times x, \frac{1}{2} \times y\right)$$

Step 6: Extract Keypoints

Step 7: if Keypoints > T_r **then**

Return Sign;

else

No Sign Found;

end else

end if

Step 8: Generate Speech Synthesis;

Step 9: End

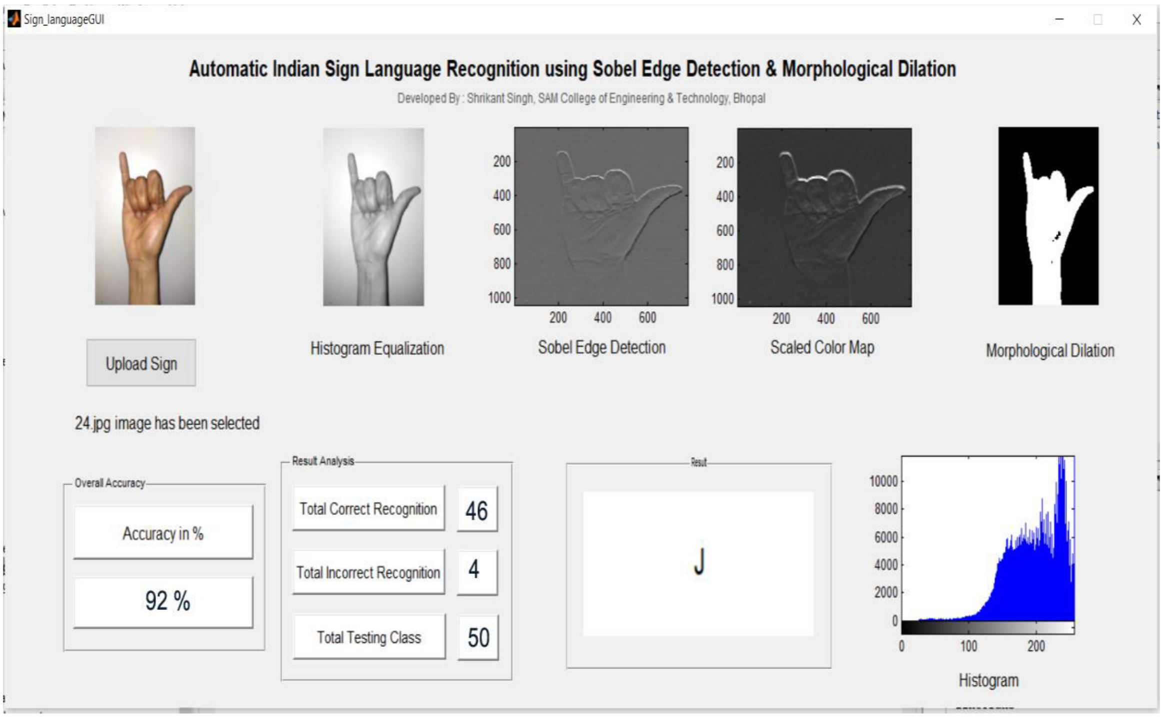


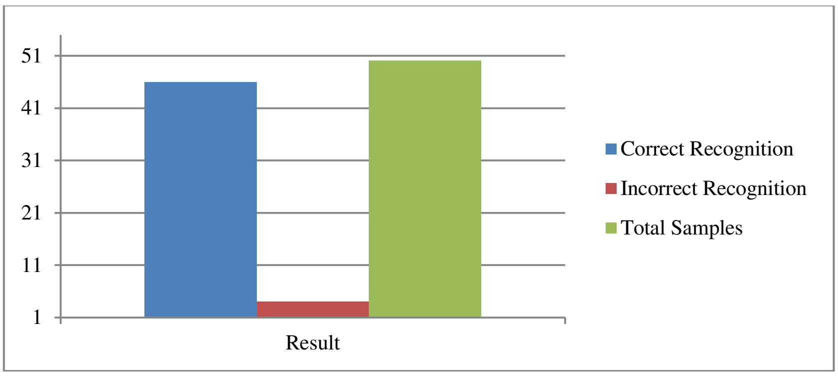
Fig. 8. Graphical User Interface

In the very first step; a sign image will acquire as 2 dimensional input array. Then the input RGB image will get converted into grayscale image for adjusting contrast using histogram equalization. $cdf_x\ (i)\ =\ \sum_j^0 P_x(j)$where cdf is cumulative distribution function, x is grayscale image, i is gray levels and P is probability. Then apply Sobel using gradient mask Gx & Gy, where Gx is horizontal mask & Gy as vertical mask and G as absolute gradient magnitude. Apply dilation for small holes filling and proper hand gesture extraction. Once the preprocessing has been done; the extracted key points will be compared with the threshold value. Once the obtained value is greater than the threshold value; it will return the result for the same.

V. Simulation Results

The simulation studies involve the various trails with distinct Indian sign language gestures. There are total number of 50 trails where 46 trails recorded as correct recognition and 4 as incorrect that may includes true positive, true negative, false positive and false negative. True positive means that there are certain trails that positively detected which returns correct recognition and few images that may contain valid sign but system is not able to detect; that entertained in the category of true negative. Similarly as false negative invalid sign detected as any template, whereas false positive means invalid sign rejected positively. So, by observing all these datasets, the perceived accuracy is 92.00 %.

Graph 1 Result Analysis



Graph 1 shows the correct recognition, incorrect recognition and total number of samples tested by the system.

Table 1 Result Comparison

	Accuracy %
B. Gupta [3]	89.50
G.Anantha Rao [4]	88.50
H. Muthu Mariappan [6]	75.00
K. Revanth [7]	90.54
S. Hayani [9]	90.02
Proposed	**92.00**

VI. CONCLUSION AND FUTURE WORK

Automatic Indian sign language recognition is a trending area where disabled people can communicate with others and remove the barriers between us. It has been implemented using sobel edge detection and morphological dilation. Most of the system uses machine learning methods to train the system with various samples for a single hand gesture. But a large dataset can consume the large amount of memory that increases the execution time where it is very important to communicate as earlier as possible with high level of accuracy. Here the system trained with various samples and able to recognize correct sign with relative features. System acquired 92.00 % of accuracy with less error rate or incorrectly recognition. In future; accuracy can be enhanced with various image enhancement technique and feature extraction method that later can be implemented for removing communication barriers.

REFERENCES

1. https://www.pinterest.com.au/pin/491103534343877366/
2. https://www.semanticscholar.org/paper/An-Automated-System-for-Indian-Sign-Language-Kaur-Gill/ebcceb337be93d44e3d6635ad5964e1e8bfaeb2c
3. B. Gupta, P. Shukla and A. Mittal, "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion," 2016 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2016, pp. 1-5, doi: 10.1109/ICCCI.2016.7479951.
4. G. A. Rao, K. Syamala, P. V. V. Kishore and A. S. C. S. Sastry, "Deep convolutional neural networks for sign language recognition," 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES), Vijayawada, 2018, pp. 194-197, doi: 10.1109/SPACES.2018.8316344.
5. Daware, Snehal & Kowdiki, Manisha. (2018). Morphological Based Dynamic Hand Gesture Recognition for Indian Sign Language. 343-346. 10.1109/ICIRCA.2018.8597417.
6. H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.
7. K. Revanth and N. S. M. Raja, "Comprehensive SVM based Indian Sign Language Recognition," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2019, pp. 1-4, doi: 10.1109/ICSCAN.2019.8878787.
8. 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, doi: 10.1109/ICCSP.2019.8698006.
9. S. Hayani, M. Benaddy, O. El Meslouhi and M. Kardouchi, "Arab Sign language Recognition with Convolutional Neural Networks," 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), Agadir, Morocco, 2019, pp. 1-4, doi: 10.1109/ICCSRE.2019.8807586.
10. A. Subhash Chand, A. S. Jalal and R. Kumar Tripathi, "A survey on manual and non-manual sign language recognition for isolated and continuous sign," Applied Pattern Recognition, 2016.
11. https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm
12. H. Cooper, B. Holt and R. Bowden, "Sign language recognition," Springer, 2011.
13. S. Ong and S. Ranganath, chez IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005.

14. M. Al-Rousan and M. Hussain, “Automatic recognition of Arabic sign language finger spelling,” International Journal of Computers and Their Applications, 2001.
15. K. Assaleh and M. Al-Rousan, “Recognition of Arabic sign language alphabet using polynomial classifiers,” EURASIP Journal on Applied Signal Processing, 2005.
16. O. Al-Jarrah and A. Halawani, “Recognition of gestures in Arabic sign language using neuro-fuzzy systems,” Atificial Intelligence, 2001.
17. T. Shanableh, K. Assaleh and M. Al-Rousan, “Spatio-temporal feature-extraction techniques for isolated gesture recognition in Arabic sign language,” IEEE Transactions on Systems, Man, and Cybernetics Part B (Cybernetics), 2007.
18. M. AL-Rousan, K. Assaleh and A. Tala’a, “Video-based signerindependent Arabic sign language recognition using hidden Markov models,” ELSEVIER, 2009.
19. M. Maraqa, F. Al-Zboun , M. Dhyabat and. R. Abu Zitar, “Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks,” Intelligent Learning Systems and Applications, 2012
20. R. Alzohairi, R.Alghonaim, W.Alshehri, S.Aloqeely, M.Alzaidan and O.Bchir, ”Image based Arabic Sign Language Recognition System,” International Journal of Advanced Computer Science and Applications, 2018.
21. Y. LeCun, L. Bottoux, Y. Bengio and P. Haffner, “Gradient-Based Learning Applied to Document Recognition,” IEEE, November 1998.
22. A. Krizhevsky, I. Sutskever and G. Hintton, “Imagenet classification with deep convolutional neural networks,” chez Proceedings of the 25th international conference on neural information processing systems, 2012.
23. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint, 2014.
24. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinov, C. Hill and A. Arbor, “Going Deeper with Convolutions,” IEEE Xplore, 2015.
25. K. He, X. Zhang, S. Ren and J. Sun, “Deep Residual Learning for Image Recognition,” 2015.
26. D. Coomans and D. Massart, “Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1. k-Nearestneighbour classification by using alternative voting rules ,” AnalyticaChimicaActa 136, 15-27, 1982.
27. C. Williams and M. Seeger, “Using the Nyström method to speed up kernel machines ,” 2001.
28. Y-W. Chang, C.-J. Hsieh, K-W. Chang, M. Ringgaard and C-J. Lin, “Training and testing low-degree polynomial data mappings via linear SVM ,” Journal of Machine Learning Research, april 2010.
29. J-P. Vert, K. Tsuda and B. Schölkopf, “A primer on kernel methods ,” Kernelmethods in computationalbiology, 47, 35-70, 2004.