Smart City Based Mobile Application for Seamless Communication of Differently-Abled

Poovam Raj T. T[1], UdhayaKumar S[2] and *Silviya Nancy J[3]

Department of Computer Science and Engineering, Rajalakshmi Engineering College, Thandalam, Chennai, Tamilnadu, India. [1]poovamraj@gmail.com, [2]mailtoudhay@gmail.com,[3]jsilviyanancy@gmail.com*

Abstract

The proposed research based mobile-application development is to aid the needs of the differently-abled (deaf and dumb) community. This application inherits the core design of artificial intelligence that eliminates the language barrier faced by the differently – abled people. The progressive features imbibed in this application would craft a revolutionary impact in their lives which would carter their needs cleverly. The base work of this application is to translate the static sign language into text that is understood by all. The application is implemented in Android operating system platform, which has the mode for wider usability of the application. The primary goal of this work is to detect and recognize the hand signs of deaf and dumb people trying to communicate and convert that into understandable text. The detection is a dynamic process since it is based on color and it can be varied in-line to the lighting conditions of the user. On the identification of the sign displayed by hand, it will be recognized by initiating the mobile application preprocessing techniques. The preprocessing technique involves attributes like conversion to HSV, and then to a binary image. As a result, from the processed image, features like the area of detecting region, bounding rectangle, min and max axis are extracted and managed, that would help in recognition of signs. Then, the related sign text will be displayed on the screen, which can be understood by the normal user. The application is also embedded with speech-to-text conversion, where the phrase spoken by the normal user will be displayed on the mobile screen as a text understandable by the differently-abled. This, in turn, makes the conversion fluent and efficient between differently-abled and normal people without the use of any specified sensor or recommended hardware model. This application can be implemented in the railway stations so that when the user enters the administer controller of this application will be available. The user can express the signs to the application which is managed by the controller and simultaneously responded with the text. By employing this in the public places, this serves under the concept of 'Smart City' in improving the lives of differently-abled.

Keywords: Artificial Intelligence; Hand sign recognition; Hand detection; Sign language–English

1. Introduction

It is observed that differently- abled people face a lot of difficulties in public places in day-to-day activities of life. The survey of World Health Organization and World Federation of Deaf has revealed the survey that there are around 72 million deaf and dumb in the world. To assist them, the government thought provokingly provides various design and infrastructures for differently-abled. For example, people with impaired vision can be guided by announcements. The deaf and dumb use hand gestures/sign for communication, so they can be provided backing with information boards and layout diagrams. The efficient practices and decisions taken by the government, smart applications are established in collaboration with smart city idea.

With the advancements in information, communication and technology have paved way for the fast spreading of "Smart City" ideas. The purpose of Smart City is to enhance the quality of the people as well as to aid the needs of different-abled in their personal and public lives more efficiently. The urge for smart concepts is to serve the human-kind with digital and information technologies to provide a smarter platform to increase the opportunities that would promote the daily needs readily with smarter solutions.

The smart city applications tremendously advocated with capabilities like the ability to monitor and control devices remotely. With the advancements in the massive production of data and ubiquitous computing, the Internet of Things (IoT) offers new insights for maintaining such infrastructures. It supports the transformation of human living by constructing more applications for public welfare. The recent progression in the development of smart devices and its application drives the entire world for all the needs. The users are readily equipped with a smart application that caters the basic needs of all the people. This has raised the use of smart phones. In recent years, there is a massive upgrade in the promising applications like wireless access, music, maps and much more. The current survey estimates that the smart phone usage would increase up to 2.5 million within 2018. The ideas of increase in smart phone usage and 'Smart City Mission', had led to the development of a mobile application that aid the deaf and dumb to converse with the public with hand gestures' that is converted to text and to speech.

In this paper, we concentrate on deaf and dumb and the obstacle they face to converse with the society to attain their needs with a mobile application. Here we have applied the scenario of deaf and dumb in public places like railway stations. The Internet of Things (IoT) along with the 'Smart City' foundations aids for the development of this smartest mobile application, where the deaf and dumb are handled by the controller with the application in a railway station.

The main concentration and goal of the application are to reduce the language barrier between the deaf/dumb community and the rest of the others. It is found that most of the time normal people find it complicated to identify what the deaf or dumb person is trying to communicate. [1] To eradicate the gap, this smart application provides a platform for communication with less complication. As it would be present in the mobile phone it would be handy to use at any place or time. The application is embedded with added concepts of image processing and detection techniques for conversation. The hand gesture (input) is captured through the mobile camera and processed through algorithms to detect the hand signs and recognize its pattern. The recognized pattern again is converted into Standard English text which is understandable by common people. Other commoners can also communicate with the disabled through the same application by recoding the voice which is fed into mobile and again it is converted in to text which can be understood by disabled also. To settle this challenging situation of deaf and dumb to live like a normal person, a smart assistant (*i.e.*, mobile application) acts as a portable electronic master to provide a standard lifestyle for the disabled which would enhance the quality of their life which is aptly depicted in Figure 1.

The architecture, explains the communication between the differently-abled and the controller of application in the public place as follows,

a) Differently-abled person nearing the help-desk with mobile application for assistance.
b) The sign is captured by the mobile camera.
c) And it is displayed as text (Hi) to controller of application.
d) The controller responds with the mobile mic (Hello, How can I help You?)
e) It is displayed to the differently-abled user.
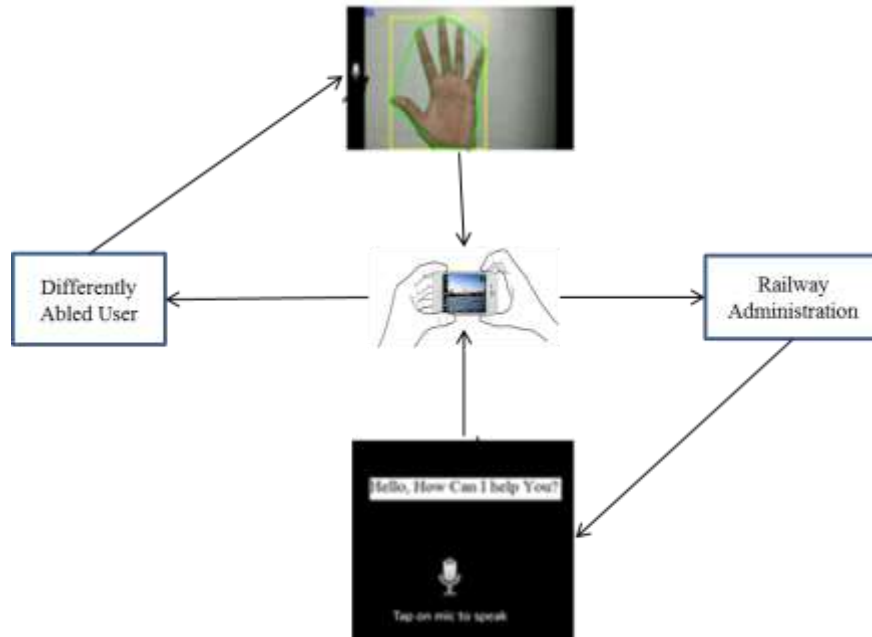
Figure 1. Smart Application for Communication of the Differently-Abled

2. Review of Literature

The mainstay of this section of the literature review is to emphasize the need for the different application to aid disable-community. The recent researchers are continuously focused on the development of a better tool or an interface application to carter to the needs of the deaf and dumb community. So, with technological and informative advances the government also takes a great leap to help them out with the help of emergence of smart city applications, which is becoming prevalent in the society.

Majzoob Kamal Aldein Omer *et al.* [3] proposes the application to educate deaf and dumb with the help of the smart mobile application. This application focuses on rendering platform for learning through images, graphs, relativity theory, and explanation with appropriate signs that offers practical and graphical oriented dais for understanding the concepts of the real world in an efficient way.

G.V.S. Subhaashini *et al.* [13] and the co-presenters of this paper proposed and developed an "Ear Hear" Android Mobile Application for differently-abled people to address normal people. The application converts the sign language into voice or text. The speech-to-sign is accomplished by Minmax technology that converts to 3D character. It also uses UTF-7 for communication purposes in day-to-day activities.

The authors, Sangeetha K and Barathi Krishna L [4] discuss the vibrant need for mobile application to support deaf and dumb community using video and image processing techniques. They have also proposed the classification of different sigh languages and procedure to identify the differences among them. The proposed application in this paper advocates equivalent text or voice output for the input hand gestures.

In the following work, the presenters Sachin Bhat *et al.* [5], declared a model with hardware devices for making the sign language, an understandable one. Here the flex sensors are placed on the hand gloves, which analyze the hand gestures according to the sign changes. Then with the help of ADC and microcontrollers, the hand signs are converted to text which is displayed to the viewers as text.

Shraddha R. Ghorpade and Surendra K. Waghamare [6], have modulated a system that paves platform for normal people to understand the sign language of deaf-mute people by using the glove to convert sign language to text and then to speech.

In the following paper, the writers, Vishal Srivastava *et al.* [7] have defined the architecture for sign language to text and then to speech conversion by handling the hand gesture images from live video using JMyron, then by using filtering algorithms the signs are detected.

Kanika Rastogi and Pankaj Bhardwaj [8], explained the usage of the smart glove in the fingers of deaf and dumb with which the sign language is converted into text.

In this particular article the proposers, Archana S. Ghotkar and Dr. Gajanan K. Kharate [9], focus on the history and interpretation of sign language and its attributes in different angles for vision based gesture classification and identification. They dominantly map the image processing techniques for conversion with the literature of sign language interpretations. They also prominently exhibit the challenges and issues involved in sign language interpretations with the help of various feature extraction and segmentation techniques.

The authors, Pallavi Verma *et al.* [10], designed a gesture based device with flex sensors and microcontrollers using smart gloves to remove the barrier of a communication lapse between the normal and differently-abled deaf and dumb users.

Dalia Nashat *et al.* [11], proposed mobile application for the deaf and dumb to communicate with other normal people. In addition to that, the presenters also collaborated an idea of incorporating quizzes, pictures, animations, games, *etc.* for entertainment.

Pallavi Verma *et al.* [12], enlightens on the design on the development of and gloves for translating the sign language to text or voice input.

The review of literature exploits the need for automotive assistance to carter the needs of the differently-abled people using a variety of techniques. Though there are many existing applications have various detection and implementation, there is no central system to identify both the sign and natural English language in both computer and mobile platform.

3. Problem Identification

Communication between the differently-abled (deaf & dumb) and normal people is always a challenging task. The objective of the study is to enable seamless communication between the deaf or dumb person and a normal person as simple conversation easing the life of a deaf or dumb person and making their disability, not an issue [2].

Figure 2. Basic Idea of the Proposed System

Initially, the hand region has to be detected to extract the feature, and they will be used as variables for the algorithm that will identify the equivalent English word to the hand sign. Then, the detected region has to be converted into a binary image. The binary image that is generated from the preprocessing technique, which includes various features of the image such as the area of the contour, area of the bounding rectangle, length of the bounding rectangle, width of the bounding rectangle, convexity points, defect points which have to be related with each other to identify the semantic of the sign that is to be used by the deaf or dumb person.

4. Detailed Description

4.1. Sign Language

The sign language (*i.e*., hand gestures) is the chief medium that is used by the disabled community for communication. This involves the combination of hand shapes, orientation, and movement of hands' for expressing their thoughts and needs. The linguists consider both the normal and sign language to be a natural one because, they do have many similarities, but there are also some significant differences between them. In sign language, each pattern is distinct from the other where pattern recognition algorithm can be used to identify the words. Figure 3 depicts the different signs for the alphabets in American Sign Language (ASL).

Figure 3. Alphabets of ASL

These signs have to be distinguished from each other to recognize them. This can be recognized by the position of hand (whether kept horizontal or vertical), a number of fingers folded, *etc.*

4.2. Recognition of Hand Gestures'' Using Image Processing Technique

The images we visualize on the computer screen are nothing but, manipulation of pixels. The screen consists of three colors such as phosphors red, blue and green. The values of these colors are given and it forms various other colors also.

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & & & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

Figure 4. Pixel Information of an Image in Matrix Representation

We can represent this pixel information in the form of a matrix, where each pixel of a row and column has its information stored in the same row and column of the matrix. We can also obtain this information for each frame and process it to identify the hand region and extract variables to detect the sign. In the static method, there is a downside which makes it would be tough to identify the hand gesture since the same object can be represented in different colors based on the lighting condition of the place.

Therefore, in this paper, we optimize a dynamic way in which the color [16] value of the hand is obtained dynamically by touching the region of hand in the mobile screen so that the radius of three pixels from this co-ordinate point will be taken and the average will be measured. This average is projected as the threshold value.

Figure 5. Processing of Pixels from a Hand Gesture

4.3. Object Tracking

In the process of object tracking, the frames are taken as input video signal. The frames are taken at a rate of 30 per second and stored along with the details of comparison in pixels. Any relative change in the pixel shows the movement of the object.
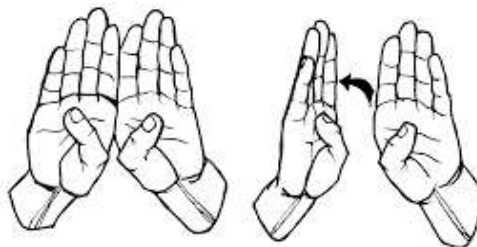

Figure 6. Object Tracking

The edges of the detected region are named as contours. The Haar cascading technique is used to locate the contours. The co-ordinates of the edges can be identified by the scanning the image by spotting only the pixel co-ordinates where the pixel is between the black and the white region. The line which is bounding the hand in red is termed as a convex hull. The bounding rectangle is the region which is identified over the region of detected image.

Figure 7. Convex Hull

The yellow dot in the below Figure 8, depicts the defect point in the images gesture. Based on the number of defected spots, the number of fingers unfolded can be recognized.

Figure 8. Convexity Defects

Before applying the threshold value, the RGB image obtained from the camera screen is converted into an HSV image which is easier for applying threshold and detection because RGB will produce more noise in the data due to the accuracy and pinpointing of the red, green and blue values. With this information, objects can be easily differentiated from each other.

Figure 9. HSV Image of the Hand

The converted RGB frame and the threshold value are applied to the HSV frame (as shown in Figure 9). At this point, the regions that fall within this threshold value are converted into white and rest of the region is converted into the black. As a result, a binary image (image only having values 0 and 1) is obtained. The binary image (as in Figure 10) is easy to process and all the variables required for the hand detection is achieved.

Figure 10. Detected Hand Region (Binary Image)

4.4. Formulation of Tables

The formulated binary image (as in Figure 9) provides the necessary variables. The variables include a minimum axis, maximum axis, contour area, bounding rectangle, the area of the bounding rectangle, points of defects and points of the convex hull. To find the bounding rectangle, the maximum points where the value 1 is found (pixel is white) have to be found and taken as a co-ordinate of the rectangle which is illustrated in Table 1.

Table 1. Finding the Bounding Rectangle

Region of maximum point	Co-ordinate point
Left maximum	X-value of top left point
Right maximum	X-value of bottom right point
Top maximum	Y-value of top left point
Bottom maximum	Y-value of bottom right point

Once the bounding rectangle, the minimum and maximum axis by the ratio of the length of the bounding rectangle to the width of the bounding rectangle is found. The contrary of the maximum axis is the minimum axis.

> If (Length of the rectangle/ Width of the rectangle >1)
>   then
>   Max_axis = X-axis;
>   else
>   Max_axis =Y-axis;

The Area ratio can be found with the bounding rectangle and the area of the detected region (area of contour).

$$\text{Area ratio} = \text{Area of bounding rectangle} / \text{Area of contour} \qquad (1)$$

The convex hull is a polygon that would engulf whole of the detected region within it. The points that form these polygons are called the convex points. These points are going to help us find how many fingers are not folded.

4.5. Detecting the Word

Using the variables obtained the word can be found. The classification of each word using the variables obtained is shown in Table 2.

Table 2. Detecting the Word Using Variables

Max axis	Area ratio	Number of Convex points	Number of Defects points	Detected word
Y	> 2	2	3 or 4	Peace
Y	> 2	1	2 or 5	Point
Y	> 2 and < 3	5	-	Hi
Y and < 0.8	-	2	1	I
Y and > 0.8	-	2	1	A
Y	< 2	5	-	B
X	> 2 and < 3	2	-	G
X	< 2	5	-	Show
X	> 3	2	-	L
X and < 0.8	-	2	1	I
X and > 0.8	-	2	1	A

Table 2 shows us how to detect the word with the sign. The letters I and A are almost similar with the variables, so the Bounding rectangle's ratio of length and width are introduced and manipulated. This shows us how even signs that are similar to each other resembling almost same values for all the variables can be distinguished from one other.
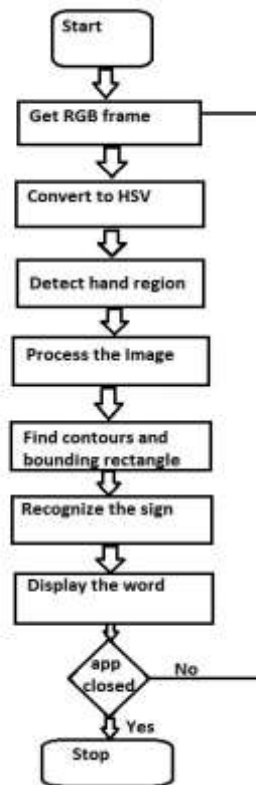
Figure 11. Flowchart Depicting the Flow of Application

5. Recommended Proposed System

5.1. Preprocessing Techniques

The block architecture explains the conversion of sign language into text. The system eliminates the barrier faced by the differently-abled while communicating. The mobile application has the following specifications. The captured image is recognized and identified by the preprocessing techniques like the conversion of RGB to HSV, then finding minimum and maximum values by detecting the bounding rectangle and contours with the region of interest. Then it is recognized and translated into respective word. The following are the modules explaining the preprocessing techniques.
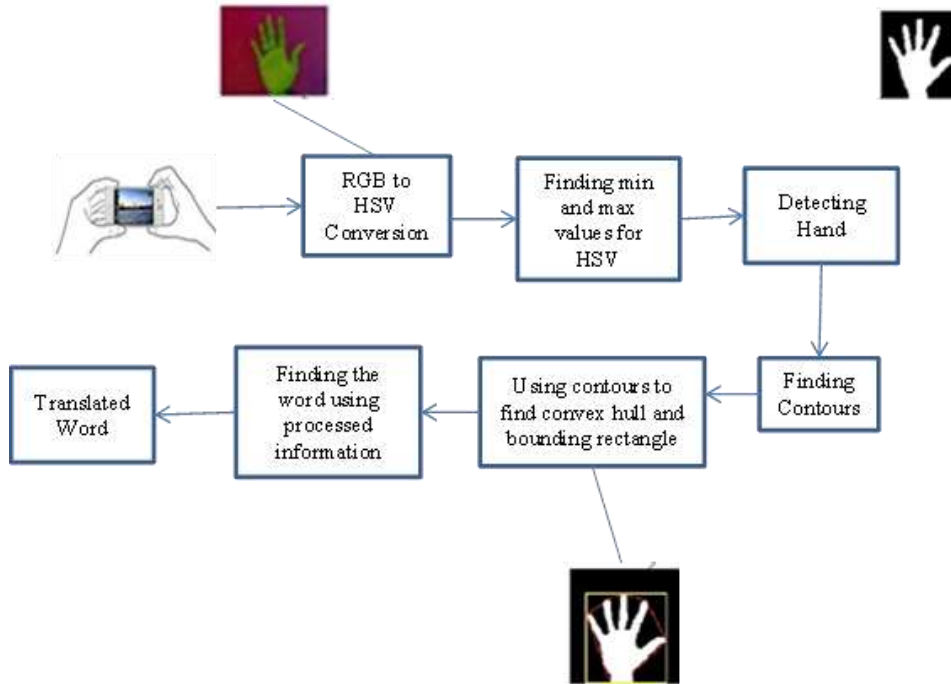
5.2. Conversion of Sign Language into Text

The modules are described as follows,
➢ Import opencv [14] into android
➢ Open camera
➢ Camera gives back RGB image
➢ RGB image is nothing but data (or) values of Red Green and blue amount present in each particular pixel
➢ Getting this pixel information as a matrix we can do operations on it
➢ The RGB image is converted to HSV.
➢ An onTouch function has to be implemented to find the region of interest (ROI).
➢ The min and max HSV value of ROI region.
  • Selected Region – 1
  • Rest of the region -0
➢ Contour, convex hull, bounding rectangle and convexity defect points are detected.
➢ Using the given information the sign should be recognized
➢ Its equivalent word should be displayed in the screen
➢ Major Axis - length and height of bounding rectangle
➢ Area of contour and convex hull - recognize the sign shown
➢ A number of contours and convexity defects - used to find whether the finger is folded or open.

5.3. Conversion of Speech to Text

➢ The user would speak into the mobile phone
➢ It would recognize and display the words spoken on the screen
➢ The screen could be shown to the deaf/dumb person

Figure 12. Preprocessing Techniques Involved For Converting Sign Language into Text

5.4. Proposed System Architecture with Results and Discussions

The proposed architecture explains the way on which the static sign language is converted into text understandable by the deaf and dumb. The application usage is demonstrated with the following scenario.

Suppose that the user (deaf and dumb), is walking into the public place (railway station), and wants certain queries to be answered. The differently-abled person normally can walk up to the help-desk to know what he wanted. And, at times the person in the help-desk may not able to understand the sign language. so, to tackle the situation and with the coherence 'smart city and Internet of Things (IoT)' interference, the mobile application can be displayed to the disable-user at the different section which can be controlled by the railway administrator and the disabled-user can be replied with the popping of text message for queries.
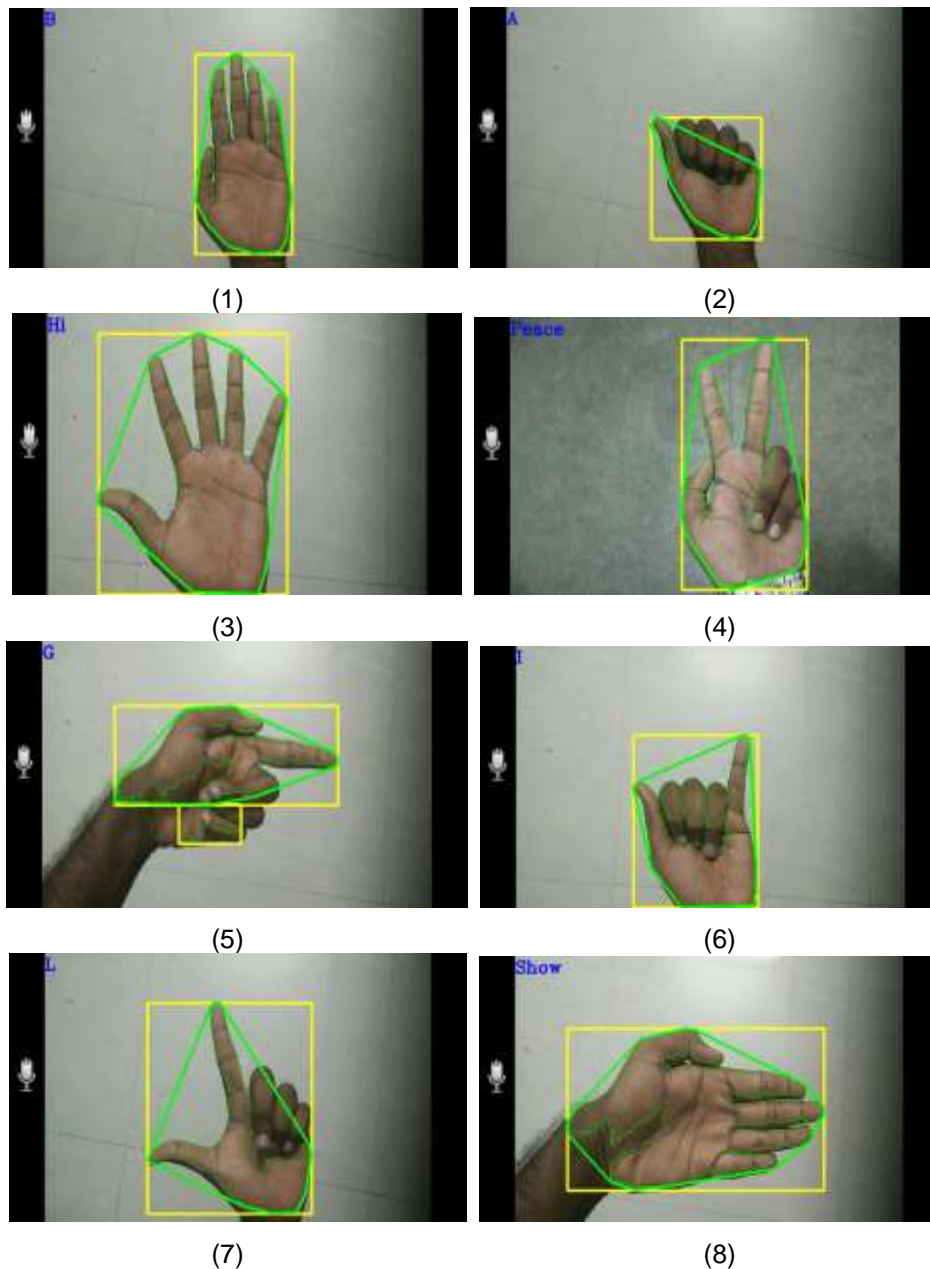
**Figure 13. Different Signs Recognized Using the Algorithm (1)"B" (2)"A"
(3)"Hi" (4)"Peace" (5) "G" (6)"I" (7)"L" (8)"Show"**

6. Conclusion

The mobile application that is developed acts as the barrier between the differently-abled and normal person by eradicating the communication barrier. The excessive movement of the hand from one place to another would be a challenging task which is not easy to solve. Because it would require a mobile phone with high end specifications (*i.e.*, 3 GB RAM) which would also result in a large set of the database. The technology with further improvement in the hardware field can be implemented in wearable like Google glass. This would make deaf or dumb people talk universally to everyone in the world. Sign language recognition is a multidisciplinary research area involving pattern recognition, computer vision, natural language processing and Psychology. As future work, all the images will be uploaded to cloud storage, and related images can be

displayed in mobile with the help of machine learning algorithm to help the user to make the conversation easier and can be used worldwide.

References

[1]   Y. Iwai, K. Watanabe, Y. Yagi and M. Yachida, "Gesture Recognition by Using Colored Gloves", IEEE International Conference on Systems, Man and Cybernetics, Beijing, China, vol. 1, **(1996)**, pp. 76-81.

[2]   N. Shimada, "Hand gesture estimation and model refinement using monocular camera—Ambiguity limitation by inequality constraints", in 3rd Conf. Face and Gesture Recognition, **(1998)**, pp. 268-273.

[3]   M. K. A. Omer, M. S. Adrees and O. E. Sheta, "Alternative Central Mobile Application Strategy to Deaf and Dumb Education in Third World Countries", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, vol. 5, no. 3, **(2015)**.

[4]   K. Sangeetha and L. B. Krishna, "Gesture Detection For Deaf And Dumb People", International Journal of Development Research, vol. 4, no. 3, **(2014)**, pp. 749-752.

[5]   S. Bhat, A. M. Ashik and C. D. Sujith, "Translating Indian Sign Language to text and voice messages using flex sensors", International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 5, **(2015)**.

[6]   S. R. Ghorpade and S. K. Waghamare, "Full Duplex communication system for Deaf and Dumb people", International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459), ISO 9001:2008 Certified Journal, vol. 5, no. 5, **(2015)**.

[7]   V. Srivastava, A. Langote and R. Barik, "Communication-Translator for Deaf Dumb and Normal People (CommuLator)", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 - 0056, vol. 3, no. 4, **(2016)**.

[8]   A. S. Ghotkar and G. K. Kharate, "Study of Vision Based Hand Gesture Recognition using Indian Sign Language", International Journal On Smart Sensing And Intelligent Systems, vol. 7, no. 1, **(2014)**.

[9]   P. Verma, S. L. Shimi and R. Priyadarshani, "Design of Communication Interpreter for Deaf and Dumb Person", International Journal of Science and Research (IJSR) ISSN (Online), **(2015)**, pp. 2319-7064.

[10] D. Nashat, A. Shoker, F. A. Swat and R. A. Ebailan, "An Android Application to Aid Uneducated Deaf-Dumb People", International Journal of Computer Science and Mobile Applications, vol. 2, no. 9, **(2014)**, pp. 1-8.

[11] P. Verm, S. L. Shimi and S. Chatterji, "Design of Smart Gloves", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, vol. 3, no. 11, **(2014)**.

[12] G. V. S. Subhaashini, S. Divya, S. DivyaSuganya and T. Vimal, "Ear Hear Android Application for Specially Abled Deaf People", International Journal of Computer Science and Engineering Communications, vol. 3, no. 3, **(2015)**, pp.1108-1114.

[13] K. Rastogi and P. Bhardwaj, "A Review Paper on Smart Glove - Converts Gestures into Speech and Text", International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169, vol. 4, no. 5, **(2016)**, pp. 92-94.