Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.199

IJCSMC, Vol. 8, Issue. 10, October 2019, pg.149 – 158

Hand Gesture Recognition Based Calculator

Sandeep Kumar[1]; Mohit Tanwar[2]; Anand Kumar[3]; Gita Rani[4]; Prashant Chamoli[5]

[1]Computer Science & Engineering Dept. & MSIT, GGSIPU, India
[2]Computer Science & Engineering Dept. & MSIT, GGSIPU, India
[3]Computer Science & Engineering Dept. & MSIT, GGSIPU, India
[4]Computer Science & Engineering Dept. & MSIT, GGSIPU, India
[5]Computer Science & Engineering Dept. & MSIT, GGSIPU, India
[1] sandeep.jaglan@msit.in; [2] mohit.tanwar1997@gmail.com; [3] anandsharma696@gmail.com;
[4] garimadeswal07@gmail.com; [5] prashantchamoli.165@gmail.com

Abstract— Gesture Recognition pertains to the process which includes understanding and interpreting meaningful movements of the body parts such as arms, face, hands or head. It allows us to talk to the machines and use them with means other than electrical devices. The technique also provides us an innovative, efficient and user-friendly way of interaction to which we are, sort of, more familiar. To convey information, the use of gestures can be quite useful in communication. This way gesture recognition plays a crucial role in our daily life. Gesture Recognition has uncountable uses, such as sign-language, machinery, robotics, computer games, virtual reality and much more. This proves that gesture recognition can be the key research area for Scientists. Hand gestures are of 2 types, static and dynamic. The three main phases of the technique on which the whole approach works are object detection, object tracking and recognizing the gesture. We developed a fast and simple calculator based on hand gesture recognition system. It uses gestures as inputs and performs calculations according to these gestures. The objective of the paper is to make the interaction between humans and computers as natural as the interaction between humans.

Keywords— Hand Gesture Recognition, Applications, Neural Networks, Gesture Recognition Issues, American Sign Language.

I. INTRODUCTION

The oldest method of communication is a gesture in the case of human beings. Gesture communication is a special non-verbal communication in which body movements/actions are used for communication. A gesture refers to an intended expression that has some meaning that involves movements of the body parts intended to convey a meaningful piece of information or to interact with the environment. The hand gesture is a splendent, intuitive and contains a good amount of message, with equal capability as languages such as written or spoken language. In gesture recognition, the gestures are read by a camera or transmitted via special gloves and various other methods are available too. The captured information is processed and used as an input to control/handle devices or applications.

Human-Computer Interaction which is also known as Man-Machine Interaction is an interrelation of humans and the computer or machine since the machine is of no use if it is not suitably utilized by a human, which makes it insignificant. The important features which should be clearly understood while designing the HCI system are usability and functionality. Usability of the system is the scope and level up to which system can do specific tasks adroitly while the functionality of the function is the set of services or the functions that the system equips to the appropriator. The system which maintains a balance between these concepts is said to be a powerful and influential performing system.

Hand gesture recognition is of 2 kinds, static gesture recognition, and dynamic gesture recognition. Static recognition refers to recognition of hand, extracting the information from the expression of the hand while the dynamic gesture recognition refers to the hand movement trajectory recognition in space and then operating on the obtained trajectory parameters. Static gesture recognition requires less computational complexity whereas dynamic is very complex but has an advantage i.e. it is very much suitable for real-time environments.

Gestures can be classified broadly into the following categories:

Head and Face Gestures

The direction of eye gaze, winking, looks of surprise, anger, fear, sadness, disgust, raising the eyebrows, nodding or head shaking, open mouth.

Body Gestures

Full-body motion involvement, as in:

1) Movement tracking of people interacting outdoors.
2) Human Gaits recognition for training and medical rehab.
3) Analysing dancer movements to generate matching graphics and music.

Arm and Hand Gestures

Sign languages, entertainment applications (virtual environment), hand poses and shapes recognition.

The gesture, what it means, can surely depend on the spatial information i.e. where it occurred, symbolic information i.e. the symbol/sign it made, affective information i.e. its emotional quality and pathic information i.e. the path it followed/took.

The main objective of this paper is to develop a calculator which takes inputs as gestures and recognize them to operate on these gestures and produce an output. The model/system will perform differently in a different environment, lighting conditions, but here we tried to minimize these limitations to as small as possible. The beauty of this system is that the user doesn't need to wear or attach any devices to use the system or interact with it.

II. APPLICATIONS

This part of the paper talks about the various applications of Hand Gesture Recognition.

A. Sign Language

It is a natural communication way of speech or hearing impaired people. It is used to interpret and explain/describe a specific subject during the communication/conversation. Various gesture recognition systems have embedded intended to interpret sign language. Generally, a camera like capturing device is used to detect and track hands, make a record of the shapes and hands trajectories and usually, they are represented by feature vectors. Then the feature vectors are matched with signs present in a grammar library to clarify whether the sign makes any sense or not.

B. Television Control

Hand gestures and postures can also be used to control the television device. Turning the TV on or off, making the volume lower or higher, turning the sound off or on, and channel activities can be easily done by using hand gesture recognition techniques.

C. Games

Computer games make use of gestures as it makes the interaction very easier from the player's point of view. To track and control the player's movement or to identify the position of the players, video games make use of gestures. A live example of this is a PlayStation in which the players use gestures to control their game.

D. Robotics

The most interesting application in this field is controlling the robots by using gestures. The wave of the hand in place of gestures can be used to control the robot. This signal should be capable of describing the desired response and the target which is the device to be controlled. In this case, the robot has a camera attached for capturing hand gestures and a gesture recognition device and a database.

III. LITERATURE REVIEW

The aim of creating a hand gesture recognition system is to create an interaction between humans and computers, where the gestures act as a spontaneous medium of conveying meaningful information [1]. The biggest challenge in the Human-Computer Interaction (HCI) systems is how to create the gestures that are well understood and interpreted by the computer [2].

Hand Gesture Recognition algorithms which are mostly used are Hidden Markov Model, which is based upon statistics [3], genetic algorithm [4] and connectionist systems [5].

HMM which is based on Statistics uses practicable relationship between the visible features procured of previous information to tackle vested uncertainty problem in processing of video, not only capable of creating the modelling of all the different dependencies of all characteristics of multiple variables of every moment, but take into account too, the probability of transition between moments, that can state the relationships regarding timings between characteristics. Though, it has to keep a specific size on the library and demand for the large quantity of computation, as larger the size of the sample library, the closer to the actual situation and the higher accuracy will be achieved in the hand gesture recognition. Also, the method has a requirement of smoothing technology to enhance the value of little probability. The genetic algorithm is used to represent the record of saved pictures using discrete quantities, controlling the discrete points and converting problems in recognizing images into problems in combinatorial optimization in discrete points. But it unable to utilize the feedback information from the network in time; furthermore, this method is troubled with long training time, slow search speed, and large training sample. An artificial neural network contains a lot of simple processing units (neurons) that are connected to form a complex information processing network; it imitates the operation of information processing, stockpile and reinstatement of the human brain neural system in a specific extent and level. On small samples, the processing adroitness is high and can easily be satisfied, but the training process requires the intervention of human and the recognition accuracy is altered by the subjective factors.

Several methods on gradient histogram have been introduced previously which used gradient and edge descriptors meant for gestures. However, they are only capable in finding gestures on background which looks simple and are not able to find hand gestures when there is mess in the background.

Various reviews recently explained gestures recognition systems and applications in our lives [9], for example, Robots, games and many surveillance [8][10]. Hence, these works show the advancement in this field, and the efforts put in to build systems with the least amount of errors using various algorithms.

The gesture recognition system is mainly divided into three steps after the input image from the camera. These are extraction method, features extraction and classification.

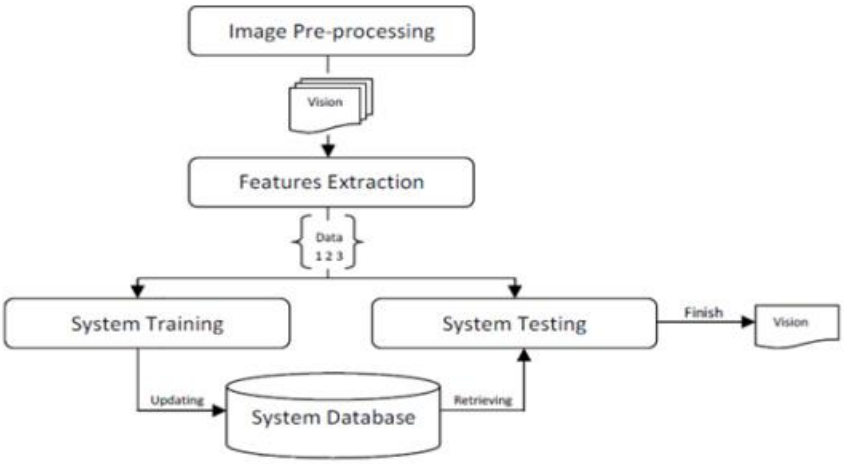


Fig. 1 Gesture Recognition System Block Diagram

E. Data Acquisition

Firstly the image is captured form a suitable input device (such as a camera) or taken an existing video as input [13].

F. Extraction Method or Image Pre-Processing

The process of recognizing the hand gesture is called segmentation or pre-image processing. This process depends on gesture – if the gesture is dynamic in nature, firstly the gesture has to correctly get discovered and then stalked and lastly segmented [11]. Firstly the hand should be discovered, a surrounding box is used for this [12], and secondly the hand has to be stalked, for tracking we have two ways; segment the video into some frames and process frame individually or use some tracking information such as skin colour, posture, shape [11].

G. Features Extraction

The successful completion of the segmentation process leads to the perfect feature extraction process [6]. Various methods can be applied for extracting the features, some methods use the shape of hands such as hand contour and silhouette, others use fingertips position, palm center, etc.[17].

H. Classification of gestures

Post analysing, molding the input method for classifying gestures is used for recognizing the gesture. The process of recognition is dependent on the properties that are selected and the algorithm applied [7]. For example, contour detection cannot be used for reorganization of gestures as many hand gestures will be generated and this can lead to misclassification [9]. The neural network has been widely used for extracting hand shape [18] and for recognizing hand gestures [19][20][21]. Other effective tools in this field are Fuzzy C-Means(FCM)[6] and genetic algorithms Gas.

Hand Gesture recognition approaches can be classified into various categories.

A. Data glove based approaches

In this approach, the user is required to wear a glove-like device, this device has a sensor that detects the moments of hands and fingers and passes information to the computer. This approach provides exact palm and finger's location and gestures [14][15].

Fig. 2 Glove Based Approach Gloves

B. Vision-based approaches

In this approach, the user requires only the camera, which captures the moments of palm and fingers (hand) for interacting between the human and computers. This approach is simple and convenient [15].

C. Colour glove based approaches:

This approach is somewhere between data glove and vision-based approach [18]. In this approach, the human wear marked gloves with some colour on them to direct the location of palm and finger for gesture recognition [16].

Fig. 3 Colour Glove Based Approach Gloves

IV. SYSTEM OVERVIEW

We divided the whole hand gesture recognition system into five subparts/modules where each module has its separate and vital role. These parts, if to named, are Setting Hand Histogram, Create Gestures, Loading Images, Training with CNN Keras, and Running Calculator. The working of these modules is described in detail below.

Fig. 4 Process of Hand Gesture Recognition

A. Setting Hand Histogram

To set the hand which is to be recognized by the system, the OpenCV library version 2 (cv2) is used to read, modify and storing back the image. For the support of multi-dimensional arrays, and matrices, the Numpy library is used. For converting a python object into a character stream to store in the file and vice versa, it is performed by using the library pickle. First of all, to capture the hand image for setting the histogram, VideoCapture class has been employed to capture image sequences. Then the image is read by the function read() which is declared in the same class. The image is then flipped vertically by using the flip() function. For colour conversion of the image, function cvtColor() is used. Then buttons for configuring and saving the image are set. When the user chooses to configure the image for the first time, then the image is captured, read and cropped into smaller images. These cropped images are then converted from Blue Green Red to Hue, Saturation, and Value model. After that, these images are then analysed by calcHist() function and generates a histogram for the images. Histogram range values are then normalized and the back projection of the histogram is calculated. The kernel used is ellipse which is obtained by getStructuingElement() which is convolved with the back projection histogram by filter2D(). The resultant image is then blurred using Gaussian and Median filter. The thresholding of the image is done to get the proper hand out of frame. Then the smaller pieces of the resultant image are merged back together to form a BGR image again which is shown on the screen along with the result of bitwise and operation between the obtained and the resultant image. If the user chooses to save the image, the webcam stops capturing further images and the results stored in the histogram are used for further operation. But if the user continues to configure the image, the steps for configuring are repeated. When the image is saved, the VideoCapture class is closed and all output windows displayed on the screen are destroyed. Finally, the result, i.e. the histogram file is converted to a character stream by the function dump(), this is process is called serializing.

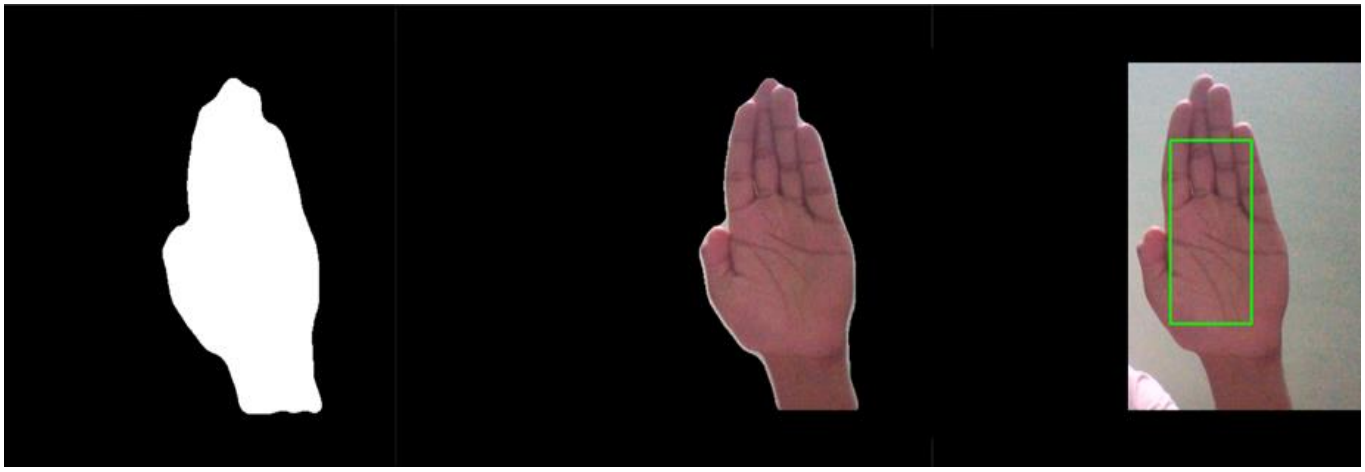


Fig. 5 Setting Hand Histogram

B. Creating Gestures

Pickle is used to serialize and deserialize a python object to save it on the disk, after converting it to a character stream this character stream contains all the information necessary to reconstruct the object in another python script. SQLite is a lightweight database that can provide a relational database management system with zero-configuration because there is no need to configure or set up anything to use it. It is a serverless database. The OpenCV library version 2 is used to modify the images. First of all, the Gestures folder and Gestures database is created, using SQL queries. Then, to store gestures in the database, the Gesture ID and Gesture Name or Text is taken as input. If Gesture ID already exists, then update the Gesture Name or do nothing. To store the gestures, if Gesture ID is 0, then it creates 1200 empty images under the folder "00", else start the camera or the webcam by using the VideoCapture class. Create the corresponding folder for the Gesture ID, and start capturing the gestures with an infinite while loop.

The frame is first read from the camera and flipped. The color conversion from BGR to HSV is done and the back-projection of the hand histogram is calculated. The image is then blurred to smoothen the edges. The thresholding of the image is done to get a proper hand out of frame.

Contours are defined as the line joining all the points along the boundary of an image that are having the same intensity. Contours come handy in shape analysis, finding the size of the object of interest, and object detection.

The image is converted from BGR to Grey and contours are extracted from the image. The boundingRect() function of OpenCV is used to draw an approximate rectangle around the binary image. This function is used mainly to highlight the region of interest after obtaining contours from an image. We can draw text on any image using puttext() method, and the pic is saved to the directory with its proper pic number.

The imshow() method of OpenCV is used to show the image on the specified window. To wait for the next keyboard event, we use the function waitKey(), where the argument is the time in milliseconds.

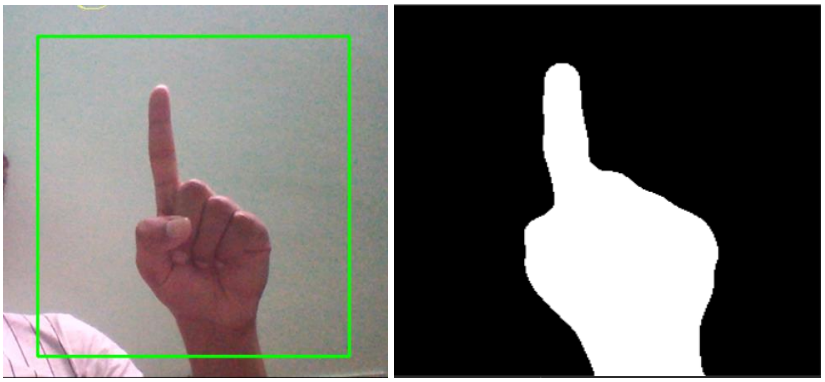


Fig. 6 Capturing Gestures

C. Loading Images

The libraries used in the module are cv2, pickle, numpy, and sklearn. The sklearn (scikit-learn) features various classification, regression and clustering algorithms.

The gesture from every folder is loaded, converted to greyscale and read by imread() function. These gestures are converted to an array by using array() function and assigning labels to every gesture. These gestures and their corresponding labels are combined and stored in a tuple.

The next step is to shuffle the gestures which are stored in the previous step by using shuffle() function. The tuple is split again and stored as images i.e. gestures in a list and their corresponding labels i.e. classes in a separate/another list.

The lists containing gestures and labels are further divided into a list of gestures which will be used to train the model, a list of gestures to test the model, a list of labels corresponding to the gestures in the training gestures and a list of labels corresponding to the gestures in the testing gestures.

These final four lists are serialized to store into the database.

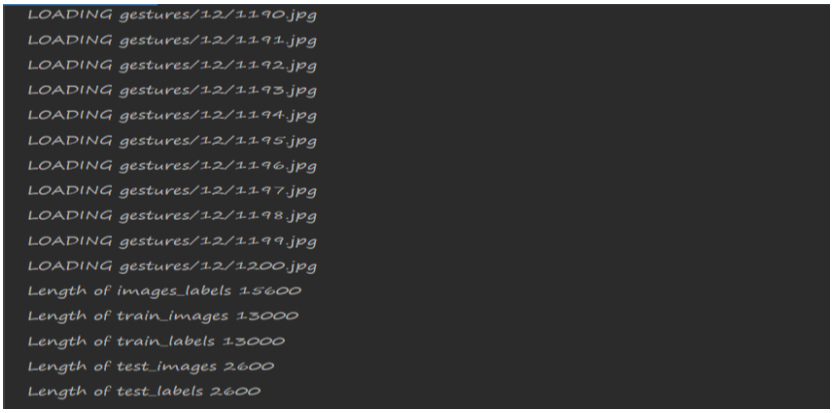


Fig. 7 Loading Images

D. Training with CNN Keras

CNN Keras is an open-source neural network library written in Python. It provides an API version of the neural network model making an easier way to create and train our model. In our case, we used sequential, Conv2D, call-backs API.

First of all, to get image shape get_image_size() function is called, this method reads the image thorough imread() method and returns the shape of the image in x and y variables.

In it, the Sequential model is used, and then the conv2D() method is used for recognizing the pattern. And these patterns are stored in the form of filters of shape 5x5.In the starting layer, we used 32 filters with activation function relu(), our model is recognizing the edges and later part, it recognizing the final shape of the object, that together helped to classify the object in the image where we have used 128 filters of shape 5x5.

To optimize our model accuracy of learning or value of parameters to maximize accuracy, we have used an optimizing algorithm known as stochastic gradient descent; it is a rectified version of gradient descent. It is a faster technique.

To prevent our model from overfitting problems, we used a state of art i.e. using dropout API.

To decide the epochs value to achieve certain level accuracy in the training part, instead of passing direct values, we have used a method of Keras library called as call-backs / checkpoints through which we can control various parameters like to stop training part after a certain point, to prevent overfitting, etc. automatically.

For pre-processing of data, to load images we have used python libraries called pickle and converted the final training labels of the data set into a categorical format using np_utilspyhton library.

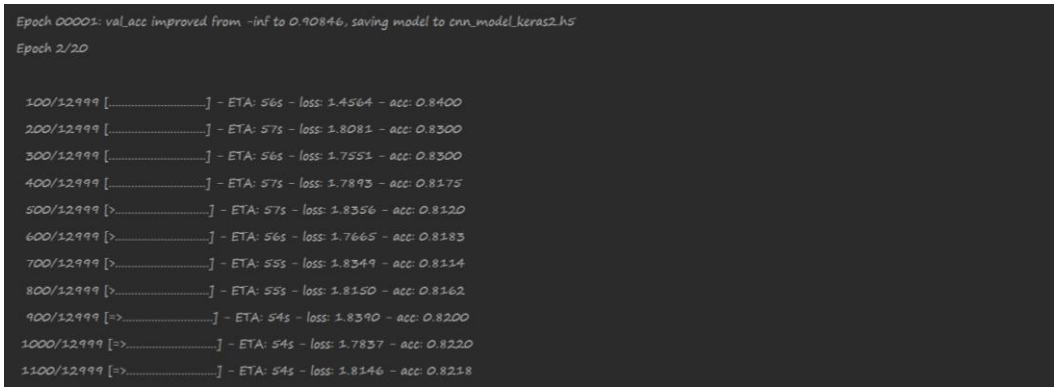


Fig. 8 Training with CNN Keras

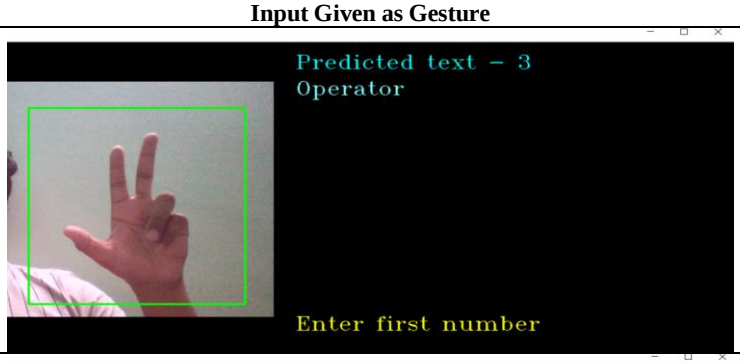
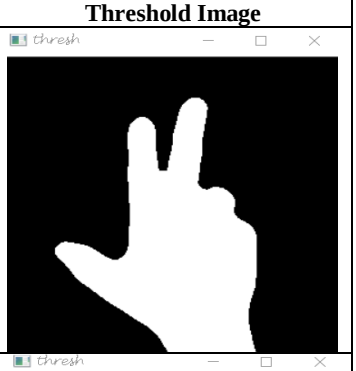
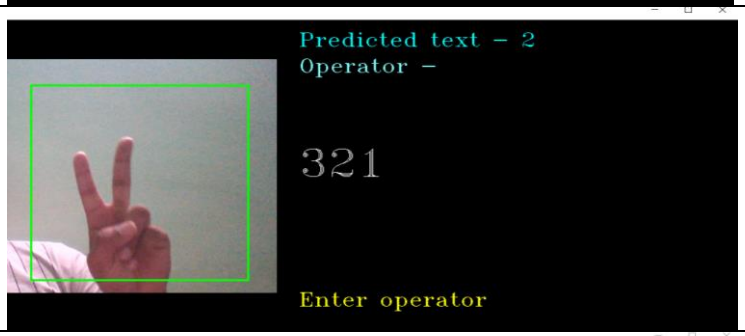
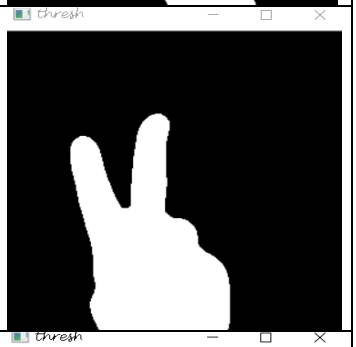
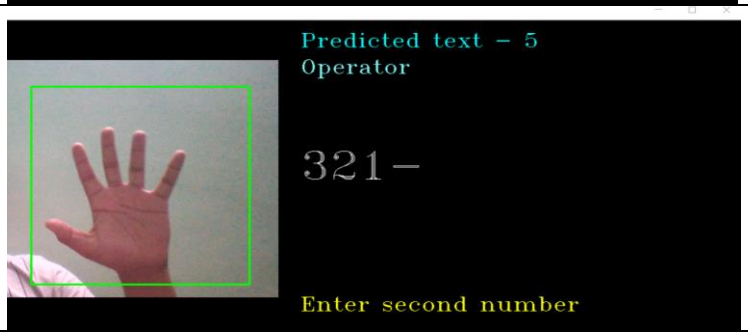
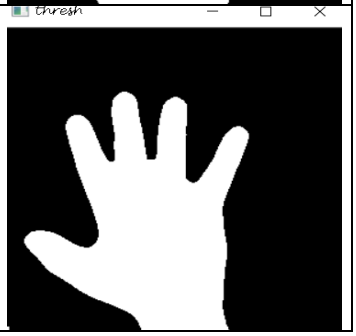
E. Running Calculator

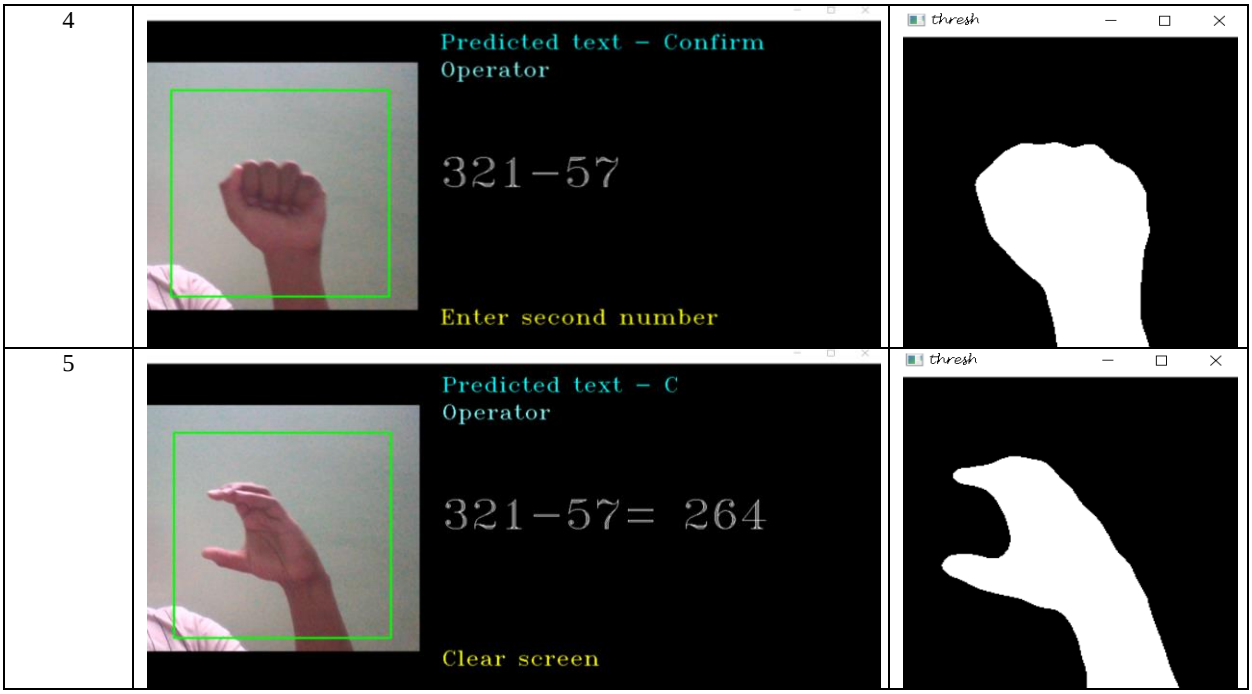
The libraries used in this module are cv2, numpy, pickle, sqlite3, and keras.

The model that was created by the CNN Keras is loaded. The webcam starts using the VideoCapture class which saves the gesture input by the user. It is converted to the standard in which it can be compared with the gestures stored in the model. The resultant gesture is passed on to the predict() function which predicts the probability of matching the gesture with stored gestures. The stored gesture with the maximum probability is selected and the label/class corresponding to the selected gesture is returned.

The histogram created by the Setting Hand Histogram is loaded. The user is asked to enter the first operand which is read and recorded by VideoCapture class and cropped into smaller pieces i.e. images. These images are then flipped vertically by using the flip() function. For color conversion of the images, we used the function cvtColor(). The back-projection of the resultant images is calculated based on the stored histogram. The kernel used is ellipse which is obtained by getStructuingElement() which is convolved with the back projection histogram by filter2D(). The resultant image is then blurred using Gaussian and Median filter. The thresholding of the image is done to get a proper hand out of frame. Then the threshold output is converted into a greyscale image. Then the image is recognized as described in the previous paragraph. To input the operator, we have assigned the different type of operations to the numerical values, for example, 1 for Addition (+), 2 for Subtraction (-). The second operand is entered in the same way as we entered the first one. Once we are done with entering the second operand, a "confirm" symbol must be given as input to the System so that the System calculates the result based on the input and display it on the screen.

TABLE I
OUTPUT

Step No.	Input Given as Gesture	Threshold Image
1		
2		
3		



V. CONCLUSIONS AND FUTURE SCOPE

This paper aimed to develop an offline Hand Gesture Recognition System. We had applied the above-mentioned process on our database of 15999 images i.e. 1200 images per character and we were able to recognize all the characters. It is determined that contour is a very important feature and can be used for discrimination between two gestures. The processing steps to classify a gesture included gesture acquisition, segmentation, morphological filtering, contour representation and classification using different techniques. The work was accomplished by training a set of feature sets which is the local contour sequence.

The area of Hand gesture-based computer-human interaction is very vast. This project recognizes hand gesture off-line so work can be done to do it for real-time purposes. Hand recognition system can be useful in many fields like robotics, computer-human interaction and so make this offline system for real-time will be future work to do. It demands a lot of research as the applications of the gesture recognition systems keep on increasing year by year. Studying and researching on hand recognition system is not only useful for improving and ameliorating the studying and working conditions of deaf peoples, but also useful and can be applied to special effects used in the film, medical research, games, robotics, and much more.

As the diversity of hand, gestures are very complex, and because of the variety of environment and background factors, the hand gesture recognition system faces many challenges.

REFERENCES

[1] G. R. S. Murthy, R. S. Jadon, *A Review of Vision Based Hand Gestures Recognition*, International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405- 410, 2009.
[2] P. Garg, N. Aggarwal and S. Sofat, *Vision Based Hand Gesture Recognition*, World Academy of Science, Engineering and Technology, Vol. 49, pp. 972-977, 2009.
[3] X. Y. Wang, G. Z. Dai, X. W. Zhang, *Recognition of Complex Dynamic Gesture Based on HMM-FNN Model*, Journal of Software, 19 (2008).
[4] X. H. Wang, H. J. Bao, *Gesture Recognition Based on Adaptive Genetic Algorithm*, Journal of Computer-aided Design & Computer Graphics, 19 (2007).
[5] F. F. Zeng, Y. Wang, G. J. Huang, *Fuzzy Neural Network Based Hand Gesture Recognition*, Mini-Micro Computer Systems, 21 (2000), 706-709.
[6] S. Mitra, and T. Acharya, *Gesture Recognition: A Survey*, IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311- 324, 2007.
[7] S. G. Wysoski, M. V. Lamar, S. Kuroyanagi, A. Iwata, *A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural*, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012.
[8] J. J. L. Junior, *A Survey of Hand Posture and Gesture Recognition Techniques and Technology*, Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA, 1999.
[9] R. Z. Khan, N. A. Ibraheem, *Survey on Gesture Recognition for Hand Image Postures*, International Journal of Computer And Information Science, Vol. 5(3), 2012.
[10] T. B. Moeslund and E. Granum, *A Survey of Computer Vision-Based Human Motion Capture*, Elsevier, Computer Vision and Image Understanding, Vol. 81, pp. 231–268, 2001.

[11] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012). *Comparative study of skin colour based segmentation techniques*, Aligarh Muslim University, A.M.U., Aligarh, India.
[12] Mahmoud E., Ayoub A., J¨org A., and Bernd M., *Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition*, World Academy of Science, Engineering and Technology 41.
[13] Surve Pranjali, Ubale V.S, *Hand gesture recognition system: A Survey*, in International journal of Inventive Engineering and Science (IJIES), ISSN:2319-9598,Volume-3,Issue-3, February 2015.
[14] Harshith.C, Karthik.R.Shastry, Manoj Ravindran,M.V.V.N.S Srikanth, Navee Lakshmikhanth, *Survey on various Gesture Recognition Techniques for Interfacing Machines based on Ambient Intelligence*, Vol.1, No.2, November 2010.
[15] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, *Vision Based Hand Gesture Recognition*, World Academy of Science, Engineering and Technology 49, pp. 972-977, 2009.
[16] Mokhtar M. Hasan, and Pramod K. Mishra, *Hand Gesture Modelling and Recognition using Geometric Features: A Review*, Canadian Journal on Image Processing and Computer Vision Vol. 3, No.1, 2012.
[17] Malima, A., Özgür, E., Çetin, M., *A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control*, IEEE 14th conference on Signal Processing and Communications Applications, pp. 1-4, 2006.
[18] E. Stergiopoulou, N. Papamarkos, *Hand gesture recognition using a neural network shape fitting technique*, Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141– 1158, 2009.
[19] W. T. Freeman and Michal R., *Orientation Histograms for Hand Gesture Recognition*, IEEE International Workshop on Automatic Face and Gesture Recognition, 1995.
[20] Min B., Yoon, H., Soh, J., Yangc, Y., & Ejima, T., *Hand Gesture Recognition Using Hidden Markov Models*, IEEE International Conference on computational cybernetics and simulation. Vol. 5, 1997.
[21] Verma, R., Dev A., *Vision based hand gesture recognition using finite state machines and fuzzy logic*, IEEE International Conference on Ultra-Modern Telecommunications & Workshops (ICUMT '09), pp. 1-6. 2009.

