Hand Gesture Controlled Robot Using OpenCV

Akhilchandran N.[1] Shejna N.M.[2]
[1]Student [2]Assistant Professor
[1,2]Department of Computer Science and Engineering
[1,2]IES College of Engineering, Thrissur, India

Abstract— Vision-based gesture recognition is an important means of human-computer interaction. This paper designs a gesture based robot system using a web-camera. After obtaining the image through the camera, various image processing techniques are performed on the gesture to remove noises, which include other operations such as binarization, erosion, dilation smoothing etc.., Then the processed gesture is given as an input to CNN model for Recognition.VGG16 architecture is used for the training CNN model.VGG16 model is loaded from keras and it is trained with our dataset and validation of model is also done. After the creation of model the input gesture captured using a web-camera which is also subjected to processing mentioned earlier. Then it is given as an input to the CNN model. The model will recognize the gesture, this recognized gesture is serially transferred to arduino board. Arduino board which in turn control the bot movement.

Keywords: Hand Gesture, Robot, Arduino, VGG16, Erosion, Keras, CNN

I. INTRODUCTION

Gestures occupy an important position in the daily communication of people. They have strong expressive force and impact, which can express information vividly, and contain a large amount of interactive information according to different gestures. Therefore, gestures are very suitable to be the media of human-computer interaction.

The communications of information between people and machines, man-machine interaction are in continuous development, but the human-computer interaction is still limited by the need of using certain physical equipment to trigger the command, and gradually cannot meet the complexity of many occasions. Therefore, to fulfill the intelligent requirements in interactive process, gesture recognition has become a hot research topic in the field of computer applications.

II. MOTIVATION

The human computer Interaction has lot of advantages of in many areas. Hand gesture controlled wheelchairs are often employed by physically challenged. Hand gesture controlled devices can be used for many military purpose. It can be used for finding blockages in manhole. it can be used to control robotic arm using human gestures. it can be also used for cleaning drainage system using a robot based on hand gesture movement.

This system can be used to identify American sign languages, it can be achieved by creating a dataset of American sign languages and training it with our created model. If identification of American sign languages is successful in future it can be used for the communication between a normal person and disability person

III. LITERATURE SURVEY

A video target detection algorithm based on pixels texture correlation background model proposed by Dongsong Yan et al. [1] The paper describes a screen target detection algorithm which is based on the background of Pixel texture information to judge models. Bayesian statistical model frame is used to predict the model. It will quickly and accurately generate reference background, accumulating less noise throughout the amount of change background, and keeping old stability. Experimental results show that compared with the Bayesian statistical model, the screen target detection algorithm which is predicated on the background of Pixel texture information to guage models has improved accuracy and reduced error rate.

Luminance and Hue primarily based approaches discriminate color and intensity data even below uneven illumination conditions. Experimental result shows the efficiency of YCbCr color space for the segmentation and detection of complexion in color images. The detection and segmentation of skin pixels victimization HSV and YCbCr color house is explained. The results of HSV and YCbCr color space based complexion detection is predicated on the choice of threshold value. These approaches discriminate color and intensity data even below uneven illumination conditions. The transformation of color pictures from RGB to HSV is time intense method

Vision-based gesture recognition proposed by Ying Wu et al. [2] The use of gesture as a natural interface serves as a motivating force for research in modeling, analyzing and recognition of gestures. which involves many interdisciplinary studies. In this, static hand posture and temporal gesture recognition is reviewed. Several application of gesture recognition are also described in the paper. This paper conclude with some thoughts about future research directions. This are using contour area and convexity defects as features for classification. It classifies the gestures using naive Bayes classifier. It have considered five hand gestures classes i.e. to show using one, two, three, four, and five fingers one by one. implemented and tested this algorithm for 10 images of each class. It gives a correct classification rate of 100. The strength of this approach includes simplicity, ease of implementation, no complex feature calculation, and no any significant amount of training or post processing required, gives higher recognition rate with less computation time. Most of the algorithm employs color based segmentation since the skin detection is more intuitive, but its main limitation is that different light condition results in change of colours very rapidly, which can cause error or maybe failures. For example due to insufficient light condition, the existence of hand area is not detected but the non-skin regions are mistaken for the hand area because of same color.

All rights reserved by *www.ijsrd.com*

Vision-based gesture recognition for human-computer interaction and mobile robot's freight ramp control proposed by Saso Koceski et al. [3] this paper presents a completely unique approach for hand gesture-based control of mobile robot's freight ramp. The research was mainly focused on solving a number of the foremost important problems that current HRI (Human-Robot Interaction) systems fight with. Presenting an easy approach to recognizing gestures through image processing techniques and one web camera, this paper address the matter of hand gestures recognition using motion detection and algorithm based on histograms.

It presents a fast, robust and accurate method for hand gestures recognition under unconstrained scenes. Experimental results show satisfactory recognition percentage of the gestures. The failure of the system to recognize the gesture is mainly due to the very changeable lighting conditions and moving objects (persons) entering the scene, operator's failure to move the hand to the proper posture. It must be emphasize that after a short experience operators get used to the system. Future work are going to be focused on algorithm improvement, by employing a combination of segmentation techniques and robot motion control by tracking the arm movement and its speed.

Image segmentation based on 2d otsu method with histogram analysis proposed by Jun Zhang et al. [4] Image segmentation plays an important role in image analysis and computer vision system. Among all segmentation techniques, the automated thresholding methods are widely used due to their advantages of straightforward implement and time saving. Otsu method is one among thresholding methods and regularly utilized in various fields. Two-dimensional (2D) Otsu method behaves well in segmenting images of low signal-to-noise than one dimensional (1D). But it gives satisfactory results only the numbers of pixels in each class are on the brink of one another. Otherwise, it gives the improper results. In this paper, 2D histogram projection is employed to correct the Otsu threshold. The 1D histograms are acquired by 2D histogram projection in x and y axes and a quick algorithm for searching the acute of the projected histogram is proposed based on the wavelet transform.

This are using contour area and convexity defects as features for classification. It classify the gestures using naive Bayes classifier. It have considered five hand gestures classes i.e. to show using one, two, three, four, and five fingers one by one. Implemented and tested this algorithm for 10 images of each class. It gives a correct classification rate of 100. Therefore, the mixture of multiple shape based features to realize gesture classification and recognition are getting a mainstream. It have presented a system which is independent of user characteristics and can be used for person independent recognition of hand gesture against uniform plain backgrounds. The segmentation of image does not affect and gives better result even if the light conditions are different. Experimental results show that the proposed method performs better than the traditional Otsu method for my renal biopsy samples. This paper introduces a 2D histogram projection analysis to unravel the matter of traditional Otsu method. Using this method, the matter of its sensitivity to the thing size are often overcome. It is very helpful for the subsequent processing and improves the success ratio of image segmentation. This are using contour area and convexity defects as features for classification. It classify the gestures using naive Bayes classifier. It have considered five hand gestures classes i.e. to show using one, two, three, four, and five fingers one by one. I implemented and tested this algorithm for 15 images of each class. It gives a correct classification rate of 100.

Design and implementation of user interface through hand movement tracking and gesture recognition proposed by Jeong Hoon Seong et al. [5] presents, how to track the movement of the hand and how to recognize the click gesture to implement a new type of user interface.Developed a wristwatch-type human–computer interface (HCI) device that can estimate and express the user's intuitive hand movements based on a 9-axis inertial measurement unit (IMU) sensor, which includes an accelerator, a magnetometer, and a gyroscope.Defined the Euler angular projection function to map the hand angle intuitively on the screen and to represent its motion reliably. It also proposed a machine-learning based gesture-recognition algorithm by extracting the window size optimized for the click gesture and collecting the accurate ground truth in a real computing environment with noise. Finally, designed a natural user interface, which is robust to the actual environment, by integrating hand motion tracking and click gesture recognition. Proved the reliability of motion by comparing the proposed hand motion-tracking function with the conventional method. In the experimental environment with noise, the press gesture recognition algorithm yielded a recognition rate of 98.94%. In conclusion, modeled the optimized click gesture-recognition algorithm and integrated the mapping functions to track hand movements, and compared the system with existing interface devices. A usability test was performed for evaluation, and usability was verified compared with existing interface equipment.

Hand gesture recognition for human computer interaction proposed by Jun Zhang et al. [6] presents A hand gesture recognition system provide a natural, innovative and modern way of non verbal communication. It has a good area of application in human computer interaction and signing. The intention of this paper is to debate a completely unique approach of hand gesture recognition supported detection of some shape based features. The setup contains one camera to capture the gesture formed by the user and take this hand image as an input to the proposed algorithm. The overall algorithm divided into four main steps, which incorporates segmentation, orientation detection, feature extraction and classification.

The proposed approach use shape base features for recognizing different hand patterns. Visually Impaired individuals will use these completely different hand gesture for writing text on electronic document like MS Office, notepad etc. Each recognized gesture is assigned with the corresponding English alphabets, numbers and some very useful key press events for generating text.

Research on the hand gesture recognition based on deep learning proposed by Jing-Hao Sun et al. [7] with the rapid development of computer vision, the demand for interaction between human and machine is becoming more

All rights reserved by *www.ijsrd.com*

and more extensive. Since hand gestures are able to express enriched information, the hand gesture recognition is widely used in robot control, intelligent furniture and other aspects. The paper realizes the segmentation of hand gestures by establishing the skin color model and Ada Boost classifier based on haar according to the particularity of skin color for hand gestures, as well as the denaturation of hand gestures with one frame of video being cut for analysis. In this regard, the human hand is segmented from the complicated background, the real-time hand gesture tracking is also realized by CamShift algorithm. This are using contour area and convexity defects as features for classification. It classify the gestures using naive Bayes classifier. It have considered five hand gestures classes i.e. to show using one, two, three, four, and five fingers one by one. Implemented and tested this algorithm for 15 images of each class. It gives a correct classification rate of 100.

Review on hand gesture based mobile control application proposed by Swapnil M.Mankar et al. [8] In existing interfaces such as Digital pen which capture only one kind of gesture and in proposed system interfaces it brings a new interaction process. It can recognize both small scale and enormous scale gestures. Surface Electro Myo Graphical (SEMG) signals are used for gesture identification. In addition, the system contains a wearable gesture sensing device and an application alongside the algorithmic framework for a mobile, are going to be designed to understand gesture-based real-time interaction then process the image. YCbCr is a color coding scheme commonly used in consumer video products such as DVD, camcorder and digital TV. Y is that the luminance component, Cb is that the blue chrominance component, and Cr is that the red chrominance component. YCbCr is a color model obtained by a linear transformation of RGB, which is computationally simple as shown in Equation 1. YCbCr separates luminance information independently, Cb and Cr channels are barely affected by light, and these three channels are separated from each other, therefore it is suitable to achieve color clustering.

The color images collected from the camera need to be binarized first to highlight the area of interest, in order to ensure the accuracy of subsequent operations. The device which is worn on the forearm, using that device user will be able to manipulate a mobile phone using predefined gestures or even personalized ones. In this system Accelerometer Sensor measure large scale gesture and SEMG sensors measure the tiny scale gesture. Both this sensors are used for the gesture recognition. The techniques which are used to recognize the hand gesture reviewed through this paper

A method for hand gesture recognition proposed by Jaya Shukla et al. [9] The paper, present a method for hand gesture recognition using Microsoft Kinect sensor. Kinect permits capturing dense, associate degreed 3 dimensional scans of an object in real time. Propose a combination of modeling and learning approach for hand gesture recognition. Use Kinect depth feature for background segmentation of hand gesture images captured with Kinect. Image processing techniques are employed to seek out contour of segmented hand images. Then calculate convex hull and convexity defects for this contour. This are using contour area and convexity defects as features for classification. It classify the gestures using naive Bayes classifier have considered five hand gestures classes i.e. to show using one, two, three, four, and five fingers one by one. Implemented and tested this algorithm for 15 images of each class. It gives a correct classification rate of 100.The paper have taken hand images with Microsoft Kinect showing gesture one, two, three, four and five fingers one by one. are obtaining binary images after applying the depth thresholding. Using image processing algorithms Contour, convex hull and convexity defects are obtained.

Handwritten recognition based on hand gesture recognition using deterministic finite automata and fuzzy logic proposed by Mohammad Zareet al. [10] presents Hand Gesture Recognition (HGR) is one of the most interesting branches in computer vision with lots of applications in mobiles, tablets, personal computers, and interactive platforms. The aim of this technology is to efficiently relate facilitate the people's communication with gadgets. Tracking hand movement in order to understand the desired operation can play an important role in the recognition of the hand gesture. In this paper, propose a new approach for handwritten recognition in the context of hand gesture recognition. My approach relies on the hand movements by means of its edge information and the movements directions. The next contribution is also minimizing the uncertainty of the handwritten recognition using fuzzy logic and deterministic finite automata (DFA). The uncertainty of the handwritten is due to the similar movements of the hands during writing the characters that decreased it using my technique. The experimental result on real-world activities shows the success and usefulness of my approach.

IV. PROPOSED SYSTEM

The research topic of this paper is controlling a robot based on gesture recognition. By calling the camera, four kinds of dynamic gestures can be identified as "Forward", "Stop", "left" and "right", and at the same time, the system can control the movement of the robotic system. In future it can be even extended to identify the American Sign Language. Architecture diagram shown in figure 1.

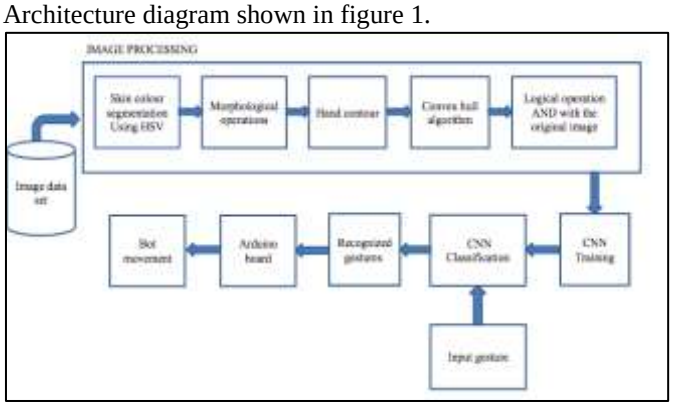

Fig. 1: Architecture diagram

The system includes five parts: gesture image acquisition, gesture image processing, CNN Training, dynamic gesture recognition and control system. In the gesture image acquisition phase, use OpenCV to call the web-camera to obtain the input image given to CNN model. In gesture image processing, performs various image processing techniques to gesture dataset obtained from

All rights reserved by *www.ijsrd.com*

kaggle.The image processing techniques performed are Skin segmentation, Morphological operations Hand contour, polygonal approximation technique. VGG16 architecture of CNN is used for CNN training.

The processed gesture dataset is given as data for training phase. and it's validation is performed with test dataset and thus created a CNN model. In gesture recognition phase the input gesture captured from web-camera is subjected to image processing, after the removal noises it is given as input to the created CNN model which will recognize the gesture Based on features extracted in training phase.

The next phase is control system, it consist of 2 sections the first one is arduino module, in this recognized gesture is serialy transferred to arduino board using pyserial library in python.arduino board is coded with the help of arduinoIDE which will produce unique action for each input gesture. The second section is robot movement, in this control the working of 2 dc motors based on input gesture for that we use L298N motor Driver.

V. IMPLEMENTATION

The hand gesture dataset used in this project is a kaggle public dataset of hand gestures. The database is composed of 10 different hand-gestures that were performed by 10 different subjects (5 men and 5 women).

Skin segmentation is the first task in image processing. In this each gesture image is converted to hsv format. We threshold the HSV image for a range of skin color. Now extract the skin color object alone. For color conversion, we use the function cv2.cvtColor(input_image, flag) where flag determines the type of conversion,for BGR $\rightarrow$ HSV, we use the flagcv2.COLOR_BGR2HSV. cv2.inRange() is used to define the range of skin color.

Two basic morphological operators are performed such as Erosion and Dilation. Erosion it erodes away the boundaries of foreground object So the thickness or size of the foreground object decreases It is useful for removing small white noises. Dilation -size of foreground object increases. Normally erosion is followed by dilation to remove noises. Because, erosion removes white noises and also shrinks our object. So we dilate it to increase object area. cv.erode() and cv.dilate() used for erosion and dilation respectively.

Contour is the boundary of an object located in an image. Contours can be explained simply as a curve joining all the continuous points having same color. The contours are a useful tool for 0bject detection and recognition. So before finding contours, apply threshold or canny edge detection.cv.threshold()used for that. To draw the contours, cv.drawContours function is used.

Convex Hull of a group of points is a tight fitting convex boundary around the points or the shape. convexity defect is a cavity in an object That means an area that do not belong to the object but located inside of its outer boundary.by removing the convexity defect from convex hull we will get only the required hand gesture.

cv2.convexHull(),cv2.convexitydefects() are used for that. at last logical operation AND is performed with the original image to ensure that all important pixels of gesture is present in the processed image.

In this project VGG16 CNN architecture is used for hand gesture recognition. which is loaded from keras and it is trained with my own dataset for feature extraction and thus developed my own customized model for gesture recognition. I will be using a sequential model. Sequential model means that all the layers of the model will be arranged in sequence. The ImageDataGenerator is used to import data with labels easily into the model.Load VGG16 pretrained model from Keras using the code. the next step is to load the dataset to the model

I used a 4 unit dense layer in the end with softmax activation as I have 4 classes to predict from the dataset. After the creation of softmax layer the model is finally prepared. Now I need to compile the model.Adam optimiser used to reach the global minima while training out model. We will also specify the learning rate of the optimiser, here in this case it is set at 0.001.model.fit_generator is used here as I am using ImageDataGenerator to pass data to the model. I will pass train and test data to fit_generator. After that model will start to train and you will start to see the training/validation accuracy and loss. after that the model will be saved for gesture recognition. then the input gesture captured using web camera is given to created model and thus identify the gesture.

Using pyserial library recognized gesture is passed serially to arduino board. The bandwidth for communication is set both in pc and Arduino as 9600 baud rate. The communication port used here is port 6..write function is used to transfer the data serially to arduino. arduino which is always in a state of listen.

Receive the data based on that issue commands to robot to move in a particular direction the working of 2 dc motors is regulated by L298N motor driver.L298N motor driver has 4 signaling pins such as 1,2,3,4 which are connected to Arduino board 7,8,9,10 pins respectively. The supply for L298N is given by 12V LI-PO batteries. when both 7 and 9 pins of arduino high then respective signaling pins 1 & 3 will be high which in turn rotates the dc motors.

If L298N terminal pin 1 is high motor 1 will rotate in clockwise direction similarly terminal pin 3 is high motor 2 will rotate in clockwise direction both these motors are connected with wheels of the robot.

Arduino output pins	Voltage High||Low	L298 N Terminal Pins	DC MOTOR M1||M2
7	High	1 ON	M1
	Low	1 OFF	
8	High	2 ON	M1
	Low	2 OFF	
9	High	3 ON	M2
	Low	3 OFF	
10	High	4 ON	M2
	Low	4 OFF	

Table 1: working of arduino

All rights reserved by *www.ijsrd.com*

If recognized gesture is forward then both 7&9 pins of arduino will be high then respective signaling pins 1 and 3 will be on thus motor 1 and motor 2 will rotate in clockwise direction which will result into forward motion. Vice-versa when 7&9 low the bot will stop. Similarly when 8&10 pin is high respective signaling pins 2&4 will be high.

When 2nd and 4th terminal pin are high m1 and m2 will rotate in anticlockwise direction respectively resulting a backward motion. Vice-versa when 8&10 low it stopes. Similarly when the motor on right side is on the robot will turns left. Vic-versa when the motor on left side is on the robot will turns right. The working left, right, stop are explained in the table above.

VI. RESULT

This paper is a combination of image processing techniques and machine learning architectures.to identify the gestures with great accuracy first of all i performed various image processing techniques on gesture image dataset to remove image background and noise often present region of interest. In this proposed system vgg16 architecture of cnn is used for the creation of model.

Which can recognize gestures. performing image processing techniques to image dataset yielded excellent results for cnn model as it removed the background and noise. The created cnn model have 98% accuracy which is better recognition rate compared to other in literatures. As this paper has high accuracy we can deploy this technique in many applications.

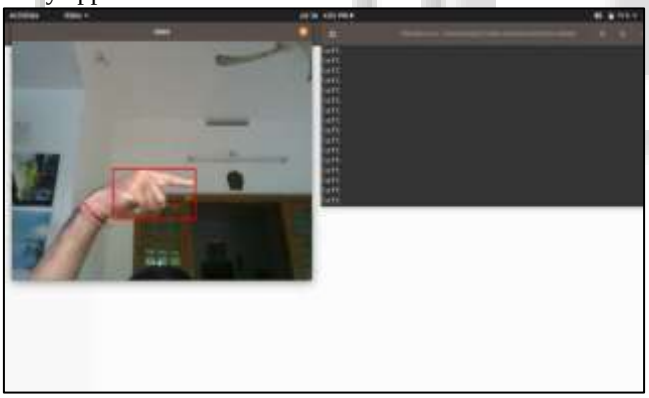

Fig. 2: Left Gesture Recognized

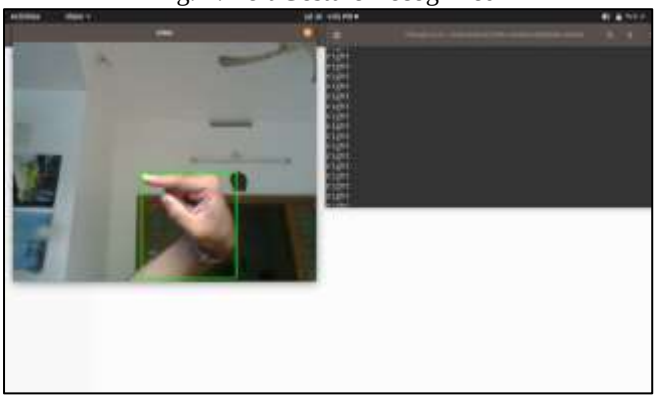

Fig. 3: Right Gesture Recognized.

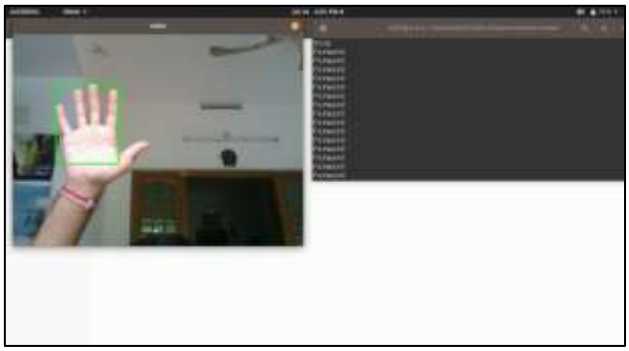

Fig. 4: Forward Gesture Recognized.

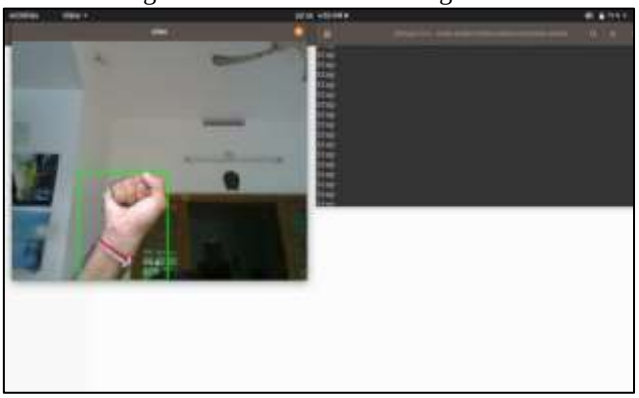

Fig. 5: Stop Gesture Recognized.

In my project the recognized gesture is used to control my robot.for this i used arduino board and L298N motor driver. Arduino board receive the gesture as a serial input.the arduino board is coded with the help of arduino IDE to receive the input and trigger the commands to L298N motor driver.

The higher recognition rate of gestures leads to greater performance. this model can identify the gestures in different environments(especially low light gesture image ,in night, low clear image etc..) The main reason behind it is the way we trained our CNNmodel. As we are using VGG16 architecture from keras we don't need to develop a model from scratch as it requires a bulk amount dataset and significant amount of time which will increase the complexity of the system as well as increase the time and cost required.

VII. CONCLUSION

Gesture movement recognition system uses a web camera to get the image, and the input image is converted from RGB space to HSV color space for object recognition. In the process of image preprocessing, morphological operation such as erosion and dilation are performed to remove the background and noises in the gesture. My project is better than many other existing systems as we are performing image processing techniques on gesture dataset and only after image processing the gesture dataset is given to Created CNN model for Training. As we are giving them only image having gestures they can easily extract features at low computational cost and in small iteration. Which not only saves our time it also reduces cost. Similarly we are using VGG 16 architecture of CNN as it showed greater accuracy of identifying gestures which is mentioned in

All rights reserved by *www.ijsrd.com*

literature survey. Our created CNN model showed greater accuracy of identifying gestures.

As a future enhancement the system can be used to identify American sign languages. This can be achieved by creating a dataset of American sign languages and training it with our created model. Instead of controlling robot we can control robotic arm and wheel chairs etc. which have many applications now days. If identification of American sign languages is successful in future it can be used for the communication between a normal person and disability person.

ACKNOWLEDGMENT

I would like to convey my heartfelt gratitude towards my guide Ms. Shejna. N.M for her constant guidance, encouraging help and inspiring words. I am thankful to the department of computer engineering for their support.

REFERENCES

[1] Jeong Hoon Seong (2018) "Design and Implementation of User Interface through Hand Movement Tracking and Gesture Recognition" IEEE Paper
[2] Yuanhao Wu (2016) "Natural Gesture Modeling and Recognition Approach Based on Joint Movements and Arm Orientations" IEEE Paper
[3] Subhankar Chattoraj (2017) "Assistive System for Physically Disabled People using Gesture Recognition IEEE 2nd International Conference on Signal and Image Processing"
[4] Mi-Young Cho (2017) "Human Gesture Recognition Performance Evaluation for Service Robots" IEEE 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI) June 28 - July 1, 2017 at Maison Glad Jeju, Jeju, Korea

All rights reserved by *www.ijsrd.com*