Care-giver alerting for bedridden patients using hand gesture recognition system

E Spandana, M Rajasekar and N Sandhya

Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Vignana Jyothi Nagar, Nizampet, Hyderabad, Telangana 500090. India.

E-mail: spandana.etrouth@gmail.com, rajasekar_m@vnrvjiet.in, sandhya_n@vnrvjiet.in

Abstract: Hand gestures are most intrinsic way of non-verbal communication. It has been difficult task for hospitals to remove the miscommunication between patients and the healthcare providers, when patients ask for help in emergency or in need. In healthcare centers like hospitals, some systems are having complicated device installation which are expensive and quite hard cumbersome. Human Computer Interaction plays a predominant role in Hand Gesture recognition (HCI). Human Computer Interaction made easy to communicate between human and machine which grabs good attention now a days. An alerting system makes a better alarm model which notifies care-givers when patients require assistance by giving defined hand gestures. The study is related to hand gesture based alerting system which uses camera module to capture the hand images. The system uses background subtraction, image recognition and speech conversion to notify the healthcare providers. It can be useful in hospitals, rehabilitation centers, or assisted living areas, where health care providers can take care of their patients more attentively. With this alerting system it increases the operational efficiency, responsiveness, patient’s and other stakeholders’ satisfaction. For recognized hand gesture text and voice is obtained as output which alerts the care-giver to help the bed-ridden patients.

Keywords: Hand gesture, Convolutional Neural Network, Image dataset generator, Machine learning, Sequential model, voice, hospital management.

1. INTRODUCTION

Humans are naturally tended to use gestures in the communication to give clear message or intentions. With today’s development of technology, communication between human-machine, the interaction between computer and human has become more recurrent. Nowadays, gesture recognition has been developed for various recognition applications. Gesture recognition is heavily encountered in human-computer interactions.

Hand gesture is a process which is used by speech and sign language; it is easily identified and understandable. Using hand gesture large amount of information is expressed. When a gesture is given by the human in the process of recognition, it is made known to the system. Hand gesture recognition has a predominant value in various applications such as sign language recognition, robot control and gesture interpretation.

2. LITERATURE SURVEY

In the wired device installation, there are many connections and limited length of wiring. Hand
gestures recognition are used in various applications, there are some papers which are discussed.

Ms. Kalpana Lamb and Prof. Mrs. Swati Madhe [1] have proposed a hand gesture recognition-based bed system where it adjusts the movement of the bed in different directions such as left, right, up and down using a DC motor and RS-232 cable for communication based on given hand gesture. The image of hand gesture is processed through background subtraction algorithm for accurate results.

Meenakshi Panwar [2] have proposed a Hand Gesture Recognition based on Shape Parameters where it detects the shape and orientation of the hand i.e., shapes based on fingers and thumb positioning, hand position according to coordinates. They use clustering concept.

Mr. Rajesh M., Ms. Bindhu K. Rajan, Ajay Roy, Almaria Thomas K., Ancy Thomas, Bincy Tharakan T, and Dinesh C [3] have developed a method which helps visually impaired people in reading the text and converting them to speech using camera module on spectacles. Use of tesseract OCR is (Optical Character Recognition) to recognize languages in text or document file converted into voice as output using e-speak algorithm.

Shining Song, Dongsong Yan, and Yongjun Xie [4] have implemented a system which uses a single USB camera to recognize the hand gesture dynamically. The camera is called by using four kinds of dynamic gestures which are “left”, “right”, “up” and “down”. Based on this gesture, system recognizes and control the movement of mouse and left-click operation is simulated.

Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang, and Yu-Bo Yuan [5] have used background subtraction method which recognizes and eliminates the hand region from background. When the hand is detected fingers and palm region is segmented. From these segmentation and finger recognition Hand gesture is recognized.

In Seop Na, Soo Hyung Kim, Chil Woo Lee, Nguyen Hai Duong [6] have implemented a method for hand gesture recognition which decomposes the hand component of finger paint dataset from human body image. In their study they identified hand gestures (finger, palm and wrist parts) which are segmented using finger paint. All the fingers are painted in different colours which make gesture recognition easy.

Vassana Inmoonnoy and Mahasak Ketchum [7] have proposed a process to detect the hand gesture from the given hand position using a webcam. The hand gesture features are extracted using different algorithms, these gestures are analysed and sent the message notification to the administrator what the patient needs.

Trong-Nguyen Nguyen, Huu-Hung Huynh [8] has implemented a method for hand gesture recognition using artificial neural network. The hand position is captured using a camera, these images are pre-processed and features are extracted using different algorithms.

Paulus Insap Santosa and Rayi Yanu Tara [9] have proposed a system of dictionary of hand gestures used in sign language. These sign language gestures are categorized as alphabets and numbers, to have a quick recognition the finger tips are marked with different colour marker caps. The dictionary is created using colour markers on each finger using Euclidean distance and nearest neighbour is used to recognize ambiguous signs.

Luigi Lamberti and Francesco Camastra [10] have implemented a method for hand gesture recognition based on coloured glove. The hand position is given to the system by wearing a coloured glove where the finger and palm parts are coloured. The hand position is given through a webcam. The hand features are extracted where it removes the background and the hand gesture. Using the distance and angle between the palm and fingers the gesture of hand is recognized.
3. PROPOSED SYSTEM

This paper represents a unique system which acts as an alerting device between healthcare providers and bed-ridden patients. The alerting system can keep the patient safe since they will be able to call for assistance when required using hand gestures. This system can be used at any healthcare providing centers to eliminate the miscommunication and to use the same in every operational way. In this section, the methodology of recognition of hand gestures was proposed and use of neural network which represents its flexibility.

The following steps will present the main parts in our method.

A. Input data

The most important step in machine learning is dataset. Usage of image dataset for hand gestures captured from a mobile phone camera of size 500x500 in a simple background. The process of identification will be less complex if the image is taken with simple background and contrast is high with the hand which has limited shadows in the image. The image dataset consists of 5 different hand gestures of 250 images each defined with our own meanings. Here are some hand gesture images used for the system refer Figure 1.

From Figure 1, index finger pointing, index and middle fingers pointing, hand gesture making C shape, little and thumb fingers pointing imply I need help, I am in emergency, Please call nurse and Go to washroom respectively.

B. Pre-processing

Pre-processing is the most significant step in the machine learning process, for the hand gesture to be identified initially the hand structure need to be detected from input data image. The hand gesture is captured with a white background with fewer shadows, where all images are in jpg format. In the first instance, the collected data is randomized and unlabeled. Five different hand gestures are considered with user-defined meanings.
C. ImageDataGenerator
Increase in size of the data gives better performance of the model because if lesser the data size, poor the performance of the model which leads to overfitting. ImageDataGenerator class in Keras library which is an iterator has been used. ImageDataGenerator dispenses the data into batches. Number of images can be requested as batch size as an argument to the ImageDataGenerator class.

D. Sequential model
Convolutional Neural Network commonly known as CNN or ConvNet is one of the algorithms in deep learning. CNN takes the input image and sets the learnable values, weights, biases to the image where the model is able to differentiate one image from another. Because of its high accuracy in speech recognition, image classification and recognition, CNN are highly used. CNN model works like a funnel modelled network which has fully connected layers with all the neurons connected and gives us a processed output. CNN has convolution layers, activation layers, pooling layers which are connected to fully connected layer which is also known as dense layer. All these connected layers, we fit particular model with training data by which weights of the neural networks defined.

In our sequential model of CNN, five Conv2D layers, five activation layer and five MaxPooling2D layers have been used. The first Conv2D layer is the input layer where we used 32 filters and 3 x 3 as kernel size. Activation layers are used to determine the final value of a neural network. ReLu is the most commonly used which gives linear positive values for all non-negative values, for negative values it gives zeros. As we are considering colour image which is combination of three colours (red, green, blue) the input value is 3, if we use grayscale image the input channel will be 1. We added 2 x 2 pool size MaxPooling2D layer which computes the maximum of each region. A batch normalization layer is added to the model which helps in increasing the speed of training data and provides stability in the model. Same as first convolution layer we add second convolution layer with 64 layers, 3 x 3 kernel size and relu as activation layer with max pooling layer followed by batch normalization layer.
We added three more layers consecutive layers. A dropout (0.2) layer is added to prevent overfitting in the model. Finally, we added flatten and dense layer with a softmax activation layer which converts the output value into a probabilistic value in the model.

From Figure 2, The flow diagram of idea implemented follows the mentioned steps in an order, starting from giving input image i.e., hand images to pre-processing. Feature extraction is done for the pre-processed image. Hand gesture images are divided into train and test sets to measure the accuracy. With the image recognition, a hand image is recognized as hand gesture and gives a word-speech conversion output.

![Flow diagram of Idea Implemented](image)

**Figure 2:** Flow diagram of Idea Implemented

4. RESULTS

We have evaluated the built model where we considered 250 images of each hand gesture images for training and 50 images of each hand gesture images for testing the model. The following table shows us the accuracy of each hand gesture. After recognition of each hand gesture, a text message displayed on UI and a voice output is given with respective to the defined hand gesture.

| Table 1: Built model evaluation |
|---------------------------------|
| **Plain Background** | **Heavy Background** |
| **Gesture** | **Accuracy** | **Speech Notification** | **Accuracy** |
| A | 94% | I need help | 50% |
| B | 84% | I am in emergency | 40% |
| C | 92% | Please call the nurse | 48% |
| D | 82% | I want to go for a walk | 42% |

*Source: Alerting system project outcome*
From the above Table 1 it is inferred that gestures A and C have higher accuracy (94% and 92% respectively) results with simple background whereas gestures B and D have lower accuracy (84% and 82% respectively) in both simple and heavy backgrounds. We can even use desired hand gestures and meanings apart from the mentioned gestures. The recognition rate of this alarming device is nearly 86.12%.

5. CONCLUSION

The implemented methodology is a real-time approach for hand gesture recognition system on alerting system for bed-ridden patients. 1200 hand gesture images i.e., 300 images for each hand gestures with defined meanings have been used. The gesture is recognized through Image segmentation and morphological operation on the captured image for further feature extraction. For recognized hand gesture text and voice is obtained as output which alerts the care-giver to help the bed-ridden patients rather than using wire-controlled system.

REFERENCES:

[1] Ms. Kalpana Lamb, and Prof.Mrs. Swati Madhe, (2016), Automatic Bed Position Control Based on Hand Gesture Recognition for Disabled Patients, International Conference on Automatic Control and Dynamic Optimization Techniques, IEEE, India.

[2] Meenakshi Panwar, (2012), Hand Gesture Recognition based on Shape Parameters, in International Conference on Computing, Communication and Application, IEEE, India.

[3] Mr. Rajesh M., Ms Bindhu K. rajan, Ajay Roy, Almaria Thomas K, Ancy Thomas, Bincy Tharakan T, and Dinesh C, (2017), Text Recognition and Face Detection aid for Visually Impaired Person using raspberry Pi, International Conference on Circuit, Power and Computing Technologies, IEEE, India.

[4] Shining Song, Dongsong Yan, Yongjun Xie, (2018), Design of control system based on hand gesture recognition, 15th international Conference on Networking, Sensing and Control, IEEE, China.

[5] Zhi-hua Chen, Jung-Tae Kim, Jinning Liang, Jing Zhang, Yu-Bo Yuan, (2014), Real-Time Hand Gesture Recognition Using Finger Segmentation, The Scientific World Journal, Volume 2014, Article ID 267672, China.

[6] In Seop Na, Soo Hyung Kim, Chil Woo Lee, Nguyen Hai Duong, (2019), Hand Component Decomposition for the Hand Gesture Recognition based on FingerPaint Dataset, Eleventh International Conference on Ubiquitous and Future Networks, IEEE, Croatia.

[7] Vassana Inmoonnoy, Mahasak Ketcham, (2017), The message notification for patient care system using hand gesture recognition, International Conference on Digital Arts, Media and Technology, IEEE, Thailand.

[8] Trong-Nguyen Nguyen, huu-Hung Huynh, (2013), Static hand Gesture Recognition using Artificial Neural Network Journal of Image and Graphics, Volume 1, No. 1, Vietnam.

[9] Paulus Insap Santosa, Rayi Yanu Tara, (2012), Dictionary of Basic Sign Language, International Conference on Future Information Technology and Management Science and Engineering, vol. 14, Indonesia.

[10] Luigi Lamberti, Francesco Camastra, (2011), Real Time Hand Gesture Recognition Using a Color Glove Springer-verlag, ICIAP, pp. 365-373, 2011.

[11] Eijaz Allibhai, (2018), Building a Convolutional Neural Network (CNN) in Keras [Online].

[12] Andrian Rosebrock, (2019), Keras ImageDataGenerator and Data Augmentation [Online].

[13] Jason Brownlee, (2019), A Gentle Introduction to the Rectified Linear Unit (ReLU) [Online].

Websites

[1] https://towardsdatascience.com/building-a-convolutional-neural-network-cnn-in-keras-329fbbadc5f5
[2] https://www.pyimagesearch.com/2019/07/08/keras-imagedatagenerator-and-data-augmentation/
[3] https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/