International Journal of All Research Education and Scientific Methods (IJARESM), ISSN: 2455-6211
Volume 9, Issue 5, May -2021, Impact Factor: 7.429, Available online at: www.ijaresm.com

Real Time Hand Gesture Recognition System

Prabhatha Nissi.G[1], Kusam Chandra Manohar Reddy[2], Govardhan.R[3], Chandu Veerendra.K[4], Aarti[5]

[1,2,3,4]Student, Department of Computer science and Engineering, Lovely Professional University, Phagwara, India
[5]Associate professor, Computer science and Engineering, Lovely Professional University, Phagwara, India

---

ABSTRACT

A hand gesture recognition device is used to communicate between a machine and a person. This paper is about Real-Time Hand Gesture Recognition (RTHGR) and to perform system control operations as designed. This application allows the use of the webcam to detect the user's hand movements and perform simple operations in response. A clear gesture must be performed by the user for the system to identify. This is captured by the webcam, which recognizes the gesture and performs the action against a set of recognized gestures. It takes a binary threshold value for gesture recognition in this procedure. For the classification process, a Convolutional Neural Network (CNN)is used. The efficiency of this method for controlling different systems will be evaluated as well as a comparison of various methods of hand recognition with the best method suitable for the task at hand being reported. This prototype can be developed in the future to assist deaf people who cannot learn operations by using a computer using this method. The project specifically focuses on gesture recognition using picture signs and perform system control operations as designed. The model comparison was made, and the best model was chosen. Thresholding is performed, followed by adaptive thresholding and neural network classification.

Keywords: ***Binary Threshold,*** **Classification, Convolutional Neural Network, Human Computer Interaction, System operations.**

INTRODUCTION

In daily lives, communication is performed via the use of vocal sounds and visual communication. However, while vocal sounds are the most useful aspect for interaction, visual communication, and facial expressions are also necessary. Including the fact that in some cases, interacting with the physical world by expressive gestures is preferable to using speech.

As we know, the vision-based technology of hand gesture recognition is a crucial part of human-computer interaction (HCI). In the last decades, the keyboard and mouse play a big role in human-computer interaction. However, due to the rapid development of hardware and software, new sorts of HCI methods are required. In particular, technologies like speech recognition and gesture recognition receive great attention within the field of HCI [1]. HCI generally includes human's face tracking, hand tracking, face recognition, hand gesture recognition, and so on [1]. Gesture Recognition determines the user intent through the recognition of the gestures or moment of the body or body parts. Hand gesture recognition has a high value in a wide range of applications such as sign language recognition, virtual reality, and sign language interpreters for disabled, robot control [2]. Especially special gameplay and Hand gesture recognition are being used in motion and performance capture systems are designed and used in industry today [3]. In the present situation, people usually do not want to touch in public areas, such as planes, touch buttons, or touch screens are used in automatic teller machines (ATM) because of hygienic considerations. In that case, hand gestures would be an excellent replacement.

Hands record a real-time video stream, which is processed for hand recognition and then trained for gesture and motion recognition. The real-time video feed through a webcam(camera) particular gesture needs to be performed by the user from the device that is connected to the devices such as laptops, computers, and phones [3]. In this project, we use machine learning techniques for training and to predict the data(info) that is hand gestures. This application performs operations according to the hand gestures detected by the webcam. To train and predict the hand gestures data the model uses machine learning techniques. The model's output can be used to control and direct various interfaces such as a computer mouse, gaming joysticks, and even a virtual simulation of the hand.

LITERATURE SURVEY

The research explores the recognition of various hand gestures and functions accordingly. Gesture recognition is an essential technique to build user-friendly interfaces. It is usable to the deaf and dumb people to make communication easy and user-friendly. It could be used in video games which allows players to communicate using a simple hand gesture. It can be used for making a touchless interface in cars like controlling the air conditioner or audio etc., which helps the driver's work easy and may save a life till some extent.

While computers have been the main reason to this success, for the most part man is the soul and center of all of this digital data. For a long time, the question has been there whether computers can be designed to analyze and acquire information from images autonomously in the same natural way that humans can. This is the province of computer vision, which is the branch of AI that ultimately aims at computers to see like a human which includes learning and being able to take actions based on visual inputs. The lack of a final scientific framework or model for human intelligence and vision from which to construct infrastructure for computer or machine learning is the issue with computer vision.

The use of photographs has a clear disadvantage. Linpu Fang et al. [7] suggested a hierarchical hand gesture recognition method based on feature covariance matrix. They used the Hidden Markov Model (HMM), the K-Nearest Neighbor (KNN), the Artificial Neural Network (ANN), the Support Vector Machine (SVM), and the pyramid Lucas-Kanade tracker for hand tracking. Local feature vector extraction for 2D dynamic HGR, Local feature vector extraction for HGR, Local feature vector extraction for HGR, Local feature vector extraction for HGR, Local feature vector extraction for HGR, Local feature vector extraction for HGR, Local feature vector extraction for skeleton based 3D dynamic HGR algorithms in the project. This is classified by capturing joints of the hand skeleton by the Intel RealSense F200 camera, This is classified by the Intel

Real Sense F200 camera capturing joints of the hand skeleton, a single low dimensional vector that is a compact dynamic hand gesture representation, a local feature extraction method based on the local motion of the hand key points, and expanding the use of the application to different systems with different data modalities, including 2D, depth-based or skeleton based 3D dynamic HGR. Centered on updated 3D convolutional neural networks, Zhi Lu et al. [6] proposed One-shot learning hand gesture recognition. SoftMax classification algorithms, Convolutional neural networks (CNN), and one-shot learning hand gesture recognition (OSLHGR) were used. Using an updated C3D network architecture, learn and extract discriminative features from different video data.

VIVA is a public dataset and multimodal fusion platform for video-based isolated gesture recognition. They improved their classification ability through continuous fine-tuning training. P. S. Neethu et al. [5] proposed a deep learning convolutional neural network-based approach for human hand gesture detection and recognition. They used algorithms such as convolutional neural networks (CNN), Nave Bayes classifier, and support vector machine (SVM). Hand ROIsegmentation with a mask image, finger segmentation, normalization of the segmented finger image, and finger recognition with a CNN classifier are used to accomplish this.

DESIGN OF SYSTEM

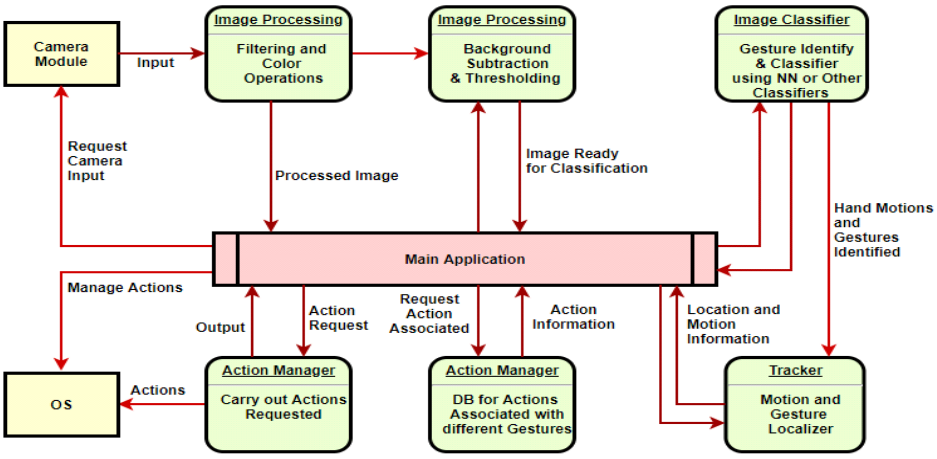
To extract features and recognize the hand gestures and act according to it following steps are followed:

1. In this paper, the human hand gestures are detected and recognized using CNN classification approach.
2. A GUI which allows the user to capture the image in real time. This phase is called image acquisition.
3. After the image is captured in real time, further step is to detect the hand.

Preprocessing steps, which are:

a) Filtering and coloring operations.
b) Background Subtraction and
c) Thresholding (Binary image conversion).

4. Image Classification using CNN classifier or only classifiers.
5. The associated gesture from the database with actions is identified.
6. Further the output gesture can be observed and system will enact according to it.



RESULTS

In this paper, python version 3.8 is used to simulate the proposed hand gesture detection and recognition technique. Open cv-python, eel, pandas, python GUI, TensorFlow, h5py, scikit-learn (0.23.2), pydot, and graphiz are all needed. The Python software is installed in windows 10 with 8 GB internal memory and executed in core i7 processor. This project was conducted in indoor and in lack of Nvidia GPU, hence the testing was conducted using only CPU.

The dataset that was used in this paper consists of 700 images for each gesture and this paper represents total nine gestures i.e., one, two, three, four, five, ok, none, palm, fist. Along with gesture recognition in this paper results were provided for system controls that was handled through gestures. The following criteria are used to determine the proposed work's overall performance, as described in this paper.

$$Recall = tp / (tp + fn)$$
$$F1\ Score = 2\,tp / (2\,tp + fp + fn)$$
$$Precision = tp / (tp + fp)$$

In above tp is true positive, fp is false positive, fn is false negative.

In Fig 1 the scalar matrix represents the classification of images according to them correctly, In total 773 images was passed to our model to predict nine gestures. The point to be considered is the 773 images were taken in random so some of images were more than 100 or even less.

```
[[ 82   0   0   0   0   0   0   0   0]
 [  0 110   0   0   0   0   0   0   0]
 [  0   0  99   0   0   0   0   0   0]
 [  0   0   0 109   0   0   0   0   0]
 [  0   0   0   0  88   0   0   0   0]
 [  0   0   0   0   0  83   0   0   0]
 [  0   0   0   0   0   0  44   0   0]
 [  0   0   0   0   0   0   0 106   0]
 [  0   0   0   0   0   0   0   0 521]]
```

Fig. 1: Scalar Matrix of our Model (CNN)

The model comparison was made, and the best model was chosen to use in this paper. The reason to use the proposed methodology is shown in Table 1 after comparing with different models and taking their Scores into view, CNN turned out to be the best model among all the models. In Table 1 model comparison between Convolutional Neural Network (CNN), Linear Regression (LR), Linear Discriminant Analysis (LDA), Decision Tree Classifier (DTC), Random Forest (RF), Gaussian Naive Bayes (GNB) were done. Some of results of Model Comparison are mentioned in Fig 2, Fig 3.

```
       Evaluating LR :
Time taken to predict 773 Images :  0  Second(/s) &  60  Millisecond
Precision: 0.640444
Recall: 0.517464
F1 score: 0.565480
[[53 10  8  5  7  7 16 33  3]
 [11 78  6 13  0  4  4  8  3]
 [ 5  8 61  9  7  9 12 13  8]
 [ 5  0  5 71  2  1  1 18  0]
 [ 0  1  3  4 67  0  0  5  1]
 [ 5  8 12  5  2 62  3 23  2]
 [ 3  5  4  2  3  0  8  6 35]
 [ 0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0]]
```

Fig. 2: Linear Regression

```
       Evaluating DTC :
Time taken to predict 773 Images :  0  Second(/s) &  105  Milli
Precision: 0.674746
Recall: 0.536869
F1 score: 0.594482
[[56  3  9  4  6  7  5 20  1]
 [ 8 85 14  8  5  6 11 24  3]
 [ 3  4 58  8  2  2  3 14  4]
 [ 5  5  3 76  5  3  4 14  4]
 [ 2  3  1  3 66  0  8 13  1]
 [ 5  7  9  7  3 65  4 12  6]
 [ 3  3  5  3  1  0  9  9 33]
 [ 0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0]]
```

Fig. 3: Decision Tree Classifier

The proposed methodology in this paper is CNN Classification from Table 1 it is clearly shown that recall score is 95.28 (recall score is also known as sensitivity (se)), F1 score is 93.46 and Precision score is (94.21) the higher these values then the methodology is considered to be best.

Table.1: Model Comparison

Testing Model	Recall Score (Sensitivity)	F1 Score	Precision Score
CNN (Used in this paper)	95.28	93.46	94.21
Linear Regression (LR)	51.7	56.5	64.04
Linear Discriminant Analysis (LDA)	51.61	56.56	65.38
Decision Tree Classifier (DTC)	53.68	59.44	67.47
Random Forest (RF)	52.22	57.59	69.75
Gaussian Naive Bayes (GNB)	12.54	16.78	31.19

The middle image shows contours along with centre as a red dot and topmost point as a blue dot. The Yellow circle is the focus region where hand is located with the red dot as its centre. The left most image shows raw input from camera with label “Region of Interest” on top and a blue rectangle which is our region of interest. The predicted gesture is shown in the top left corner of the blue rectangle, along with the neural network's confidence percentage. This confidence percentage cannot be taken as single parameter for measuring accuracy of the model rather just a parameter for confidence. The bottom centre shows ‘dX’, ‘dY’ variables followed by direction which showcase the movement of the hand placed inside the region of interest. The right most image shows the binary threshold of the raw image, where the white pixels are the hand detected and back pixels are the background, obtained after series of image processing techniques. Consider the testing prediction of the neural network for different gestures inputs like 1, 2, 3, 4, 5, ok, fist, none, palm. Examples of the prediction is shown in different backgrounds in following figures.

Fig. 4 : Test Output for input 1.

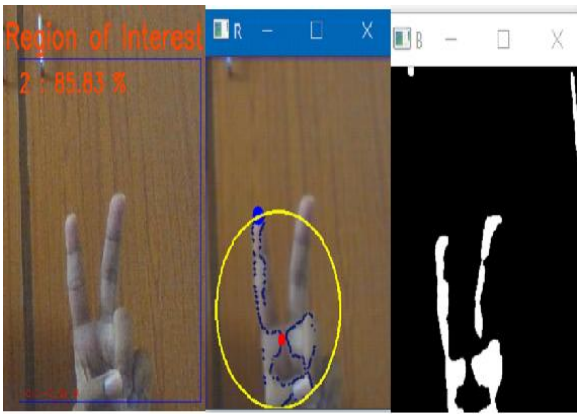


Fig. 5: Test Output for input 2.

REFERENCES

[1]. Merin Joy, D. and Rajesh, D., 2017. Hand Gesture Recognition from Surveillance Video Using Surf Feature. International Journal of Engineering Research and Science, 3(5), pp.82-89.
[2]. N U, I.R D. Development of an Efficient Hand Gesture Recognition system for human computer interaction. International Journal Of Engineering And Computer Science. 2016. doi:10.18535/ijecs/v4i12.5
[3]. Sensors and Materials, 2018. Real-time Hand Gesture Recognition System and Application. p.869
[4]. Rohith Gudelli, Madi Prashanth Reddy, Harini, Dr SP singh. "Classifying the Fingers to Recognize Hand Gestures by using Open-Source Computer Vision". IRJET (2011).
[5]. P. S. Neethu, R. Suguna, Divya Sathish. "An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks". Springer indexed Scopus (2020).
[6]. Zhi Lu,Shiyin Qin, Xiajioe Li, Lianwel Li, Dinghao Zhang. "One-shot learning hand gesture recognition based on modified 3D convolutional neural networks" Springer indexed Scopus (2018).
[7]. Linpu Fang, Guile Wu, wenxiong kang, Qiuxia Wu, Zhiyong Wang, David Dagon Feng (2018). "Feature covariance matrix-based dynamic hand gesture recognition". Springer indexed Scopus.
[8]. Xude dong, Yuanping Xu, Zhijie Xu, Jian Lu, Li lu. "A Static Hand Gesture Recognition Model based on the Improved Centroid Watershed Algorithm and a Dual-Channel CNN". IEEE (2018).
[9]. Muhammad Inayat Ullah Khan (2011). "Hand Gesture Recognition System". diva-portal.org.
[10]. A. Graves and J. Schmidhuber, "Offline handwriting recognition with multidimensional recurrent neural networks," in Advances in neural information processing systems, 2009, pp. 545–552.
[11]. A. Graves, M. Liwicki, H. Bunke, J. Schmidhuber, and S. Fernandez, "Unconstrained on-line handwriting recogni- ´ tion with recurrent neural networks," in Advances in Neural Information Processing Systems, 2008, pp. 577–584.
[12]. Graves, S. Fernandez, and J. Schmidhu- ´ ber, "Multi-dimensional recurrent neural networks," CoRR, vol. abs/0705.2011, 2007. [Online]. Available: http://arxiv.org/abs/0705.2011.
[13]. Ahmed Kadem Hamed AlSaedi, Abbas H. Hassin AlAsadi. "A new hand gestures recognition system".(2020). Indonesian Journal of Electrical Engineering and Computer Science(IJEECS).
[14]. Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang and Yu-Bo Yuan. "Real-Time Hand Gesture Recognition Using Finger Segmentation". The Scientific World Journal (2014).
[15]. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: CVPR, pp. 1725–1732 (2014).
[16]. Sharma S, Jain S, Khushboo (2019) A static hand gesture and face recognition system for blind people. In: 2019 6th international conference on signal processing and integrated networks (SPIN), Noida, India, 2019, pp 534–539.
[17]. Wang Z, Chen B, Wu J (2018) Effective inertial hand gesture recognition using particle filtering based trajectory matching. J Electric Comput Eng 2018, 6296013.
[18]. Wang Y, Jung C, Yun I, Kim J (2019) SPFEMD: super-pixel based finger earth mover's distance for hand gesture recognition. In: ICASSP 2019–2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), Brighton, United Kingdom, 2019, pp 4085–4089.
[19]. Marium A, Rao D, Crasta DR, Acharya K, D'Souza R (2017) Hand gesture recognition using webcam. Am J Intell Syst 7(3):90–94.
[20]. Rahman MH, Afrin J (2013) Article: hand gesture recognition using multiclass support vector machine. Int J Comput Appl 74(1):39–43.