IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CONTROLLING MULTIPLE APPLICATIONS WITH HAND GESTURE USING CONVOLUTION NEUTRAL NETWORK

[1]Raj Kumar E M, [2]Smitha G Adiga, [3]Md Altaf Raja, [4]T V Anuhya, [5]Dr. C Nandini

[1234]Student[BE], Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management(DSATM), Bengaluru, Karnataka, India.

[5]Vice principle, Professor and head, Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management(DSATM), Bengaluru, Karnataka, India.

Abstract--In today's world, everybody opts for quick interaction with complicated systems that guarantee a fast response. Thus, with increasing improvement in technology, interval and simple operations are the issues. Here where human-computer interaction comes into play. The vision-based technology of hand gesture recognition is vital part of human-computer interaction (HCI). Owing to the fast development of hardware and software system, new forms of HCI strategies are needed. Technologies like voice recognition and gesture recognition receive nice attention within the field of HCI. Deaf and dumb people lack in correct communication with traditional individuals and realize it troublesome to properly express themselves. Thus, they're subjected to face several problems during this regard. The sign language is extremely common among them and that they use it to convey themselves. In this project, with the assistance of computer vision and deep learning techniques, user hand movements (gestures) are employed in time period to manage multiple applications. In this project, seven gestures are outlined to manage multiple applications using hand gestures. This application allows the user to use their device camera to spot their gesture and execute the management over the multiple applications (without any extra hardware). It will increase potency and makes interaction easy for the user to manage his/her laptop/desktop from a distance.

Keywords— Convolution Neural Network, Deep Learning, Hand Gesture Recognition, Multiple Application management, OpenCV, Streamlit.

I. INTRODUCTION

Human communicate with one another by transfer their concepts, thoughts and experiences to the individuals around them. Human communicate with one another by transfer their concepts, thoughts and experiences to the individuals around them. Hand gesture is one among the strategy employed in language for non-verbal communication. In this paper, approach for deaf-people interfacing with computer vision is conveyed. The strategy is straightforward enough to supply real time recognition and works fittingly to most letters. Currently we have a tendency to perpetually hear concerning new technology that improves our style, that produces our life easier. This analysis result in new inventions and creating one's life easier. However, we've terribly less analysis for Deaf & dumb individuals. Deaf and dumb individuals perpetually notice difficulties to speak with traditional individuals. They feel to not communicate and that they never able to convey their feelings.

Gesture could be an image of physical behavior or emotional expression. It includes body gestures and hand gesture. Essentially there are 2 main language recognition approaches: image-based and sensor based. Gesture recognition is gaining importance in several applications areas like human interface, communication multimedia system and security. This technique works based on video process technology. The detection of hand gestures will be done with webcam. The images are then regenerate into normal size with the assistance of pre-processing. Currently deaf and dumb individuals have become more & outgoing, and in contrast to old days, they are not depending on anybody for communication. So, for such individuals it's vital that the overall public around them should even be able to perceive what they need to inform those using sign language.

Gesture recognition could be a technology aimed toward providing live-time knowledge to a computer to execute commands the user desires. Individuals don't need to type something with keys or to tap on the screen to perform a selected action. The device's motion detector will understand and interpret the person's movements because the primary supply of knowledge input. Therefore here we've developed its hand tracking supported by neural networks that incorporates a stack of 2 deep-learning technologies. Does one keep in mind the days once the touchscreen technology initial came out? Several people were happy to prevent using physical buttons to regulate our TVs, direction systems in our cars, smart phones and different devices switch to smart screen's technology instead. This saved users' time and was additionally convenient and pleasurable to use.

In this paper we are concerned about controlling the multiple application with hand gestures using convolution neural network. Some of the use cases include watching a movie or having to change the attention to another direction by looking or moving away from the computer, which leads to missing an interesting clip of the movie/cooking tutorial/workout videos. For all the above distractions, the user has to repeatedly pause or play and rewind the video, which eventually leads to the usage of dirty or sweaty hands to control the player.

The solution to all or any the on top of situations is that the proposed system. An web application that uses the device camera to provide users a touch-free and remote-free management over any applications while there is no need of any special hardware. The sole data required is that the mapping of gestures to controls. This application is extremely productive, makes life easier and comfy by controlling the computer device from a distance.

II. LITERATURE SURVEY

N. Krishna Chaitanya, R.Janardhan Rao [8], they presents Hand gesture recognition, especially in the area of human computer interaction, hand gesture has gained significant interest in recent years. Unlike user interfaces which uses a mouse or keyboard, interfaces with gestures provides with interactions that mirror their experience of life away from display screens. These interfaces recognise and interpret user body movements. These interactions can feel more natural and simpler for users and doesn't require special devices. As gesture commands are found to be natural and simpler for users, the development of the gesture-based system interface have become an important research area. Application dependent system which makes it complex to transfer one gesture control interface into multiple applications. The drawback is that there was an increased computational cost during training the system and recognition; for example, detecting hands in all possible configurations or heads in all possible orientations is not currently possible at interactive frame rates.

Yashas J, Shivakumar G [10] they presented the literature survey performed on Hand Gestures Recognition system. Possible methods such as sensors that are worn on hands and camera is discussed with respect to data acquisition. Cameras provide a better advantage by not limiting the physical movement of the hand but bring in new challenges like region of interest, background noise and lighting. These challenges are solved to an extent by devices like the Microsoft Kinect. Learning algorithms like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN) and Deep Learning. The usage of CNN is challenged by Adapted Convolutional Neural Networks (ADCNN) as the latter uses data augmentation to analyze data in better way to allow for better classification. This paper concludes by stating that data augmentation is a field yet to be explored in detail.

Parshwa P. Patil, Maithili J. Phatak, Suharsh S. Kale, Premjeet N. Patil, Pranav S. Harole, Prof. Dr. C. S. Shinde[3] presented ,Deaf and dumb people find it difficult to properly use digital devices due to lack in proper communication with digital devices. Thus, they were facing many issues in this regard. The sign language is very popular among them and deaf and dumb people used it to express themselves. They make use of sign languages to communicate which is hard to interpret by the devices. There is a necessity of building up a tool that interprets the gestures into text and speech. This will be a good step to make the communication possible between the deaf and dumb people and the possible tools. They are providing an interface for deaf and dumb who can't speak to interact with digital devices. Although this approach can identify a big number of gestures, it also has some drawbacks, such as missing some gestures because of the classification program accuracy contrast.

Neha S.Rokade, Harsha R. Jadhav, Sabiha A. Pathan , Uma Annamalai [7], they presented Hand gesture recognition with respect to human-machine interfaces is being developed greatly in recent days. Many visual hand gesture recognition systems operate or show successful results only in restricted background, because the disturbance of lighting and background are not plain. In order to recognize the various hand gestures, they will provide a non-complex and with speed motion history image related systems. In their system, they mainly focus on applying pointing behaviour for the human machine interface. A button has an explicit purpose and simple to find gestures, however, may be arbitrary and are usually more complex to discover.

Niranjani V, Keerthana R, Mohana Priya B, Nekalya K, Anantha Krishnan Padmanabhan[1], theypresent, the Human Computer Interaction progresses toward interfaces that seem to be natural, simple, intuitive to use rather than the customary usage of input devices such as keyboard and mouse. Because of its diversified application and the potential of interacting with machines proficiently, hand gesture recognition system is one of crucial techniques to build user friendly interfaces. Hand gestures including the movement of arms, fingers or hands are considerable for interaction. The present system for the application access is arduous and inflexible for user with hand deformity and blindness regarding the human-computer interactions. A deep convolutional neural network (DCNN) is put forward in that paper, to use hand gestures recognition and immediately classify them by preserving even the other area(not-hand area) without any detection or segmentation process. Hence the proposed objective of this paper is to use different hand gestures via integrated web camera with the help of DCNN concept being beneficial for the people with hand

disability and visually impaired people. Items in the background or distinct features of the human may make recognition more complex. The distance from the camera and camera's quality and resolution also causes variations in recognition accuracy.

III.PROPOSED SYSTEM

The proposed hand gesture detection system for multimedia application control is broadly classified into two stages. The first stage performs the detection of hand gestures and in the second stage keyboard and mouse control are integrated with each gesture.

The hand gesture recognition is implemented using OpenCV(open source computer vision) and deep learning techniques with a custom-built dataset. A new dataset is created that contains 7 gestures as shown in Fig. 1. 1.

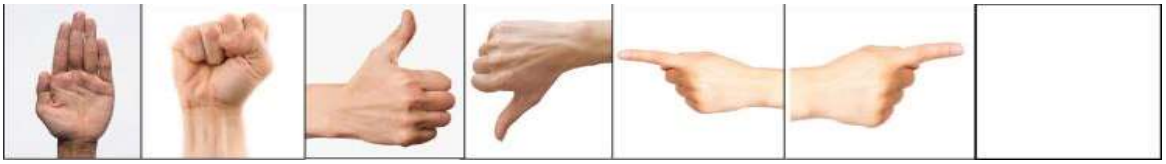

Fig. 1.2 Raw Data collected for the dataset

These raw images from the dataset are converted to black and white images as shown in Fig.1. 2.

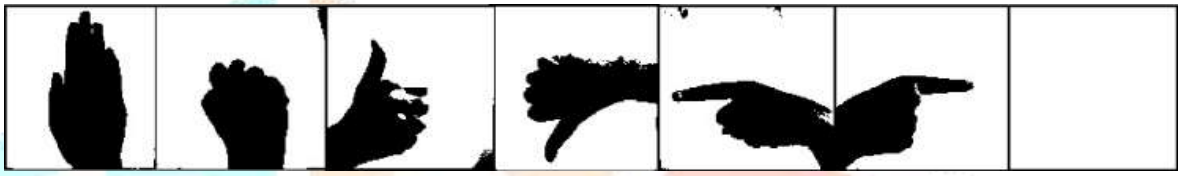

Fig. 1.2 Gesture data after preprocessing

A. *Proposed Workflow*

The system is divided into five modules. Fig. 1.3 shows the overview of the system workflow, emphasizing the gestures and their corresponding controls.

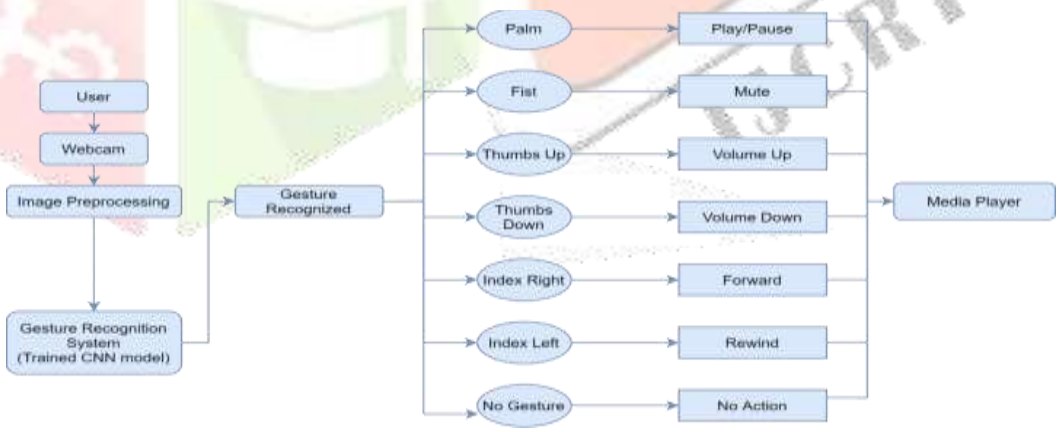

Fig. 1.3 The System Design Workflow

1) *Image Acquisition and Pre-processing*

When the user performs hand gestures in front of the webcam. The image frames are collected from the live video using the OpenCV. These images are converted into black and white images as shown in Fig. 1.4, to improve accuracy in predicting gestures and then stored in respective directories.

While the camera is on, two frames are displayed on the screen and the user can capture images frame-by-frame using the read function and the mirror image is simulated. The user has to place the hand in the Region Of Interest (ROI) i.e., the bounding box, and perform the gestures. The frames are extracted from ROI and resized to 120x120x1. The count of the number of images in each directory is printed onto the screen. The count is increased every time user captures an image by pressing 0 to 6 number keys on number keys on the keyboard and the images are saved to their respective class directory.

The preprocessed image can be seen in the small frame while capturing and these images will be stored in the dataset. The user can exit after data collection by pressing the escape key on Keyboard.

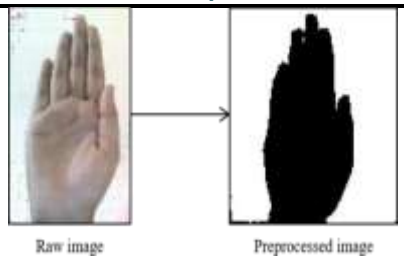

Fig. 1.4 Data Preprocessing

2) *Feature Extraction*

Import Keras models and hidden layers required to build convolutional networks. The CNN model is built using a hidden input layer followed by two convolution layers. An activation operation called ReLU and a pooling layer called MaxPooling are added after every convolution layer. A flattening layer is added along with two fully connected layers, one with RELU activation and another with softmax activation which is used for classification.Lastly, the compilation of the CNN model is performed by implementing adam algorithm as an optimizer, categorical_crossentropy as a loss function method to find error and accuracy as the performance metrics to evaluate the model.

3) *Train the model*

Next, ImageDataGenerator class is used to generate batches of images to train and validate the model. The fit function is used to train the model with a fixed number of epochs. After training, the trained model is saved in JSON format and the weights are saved directly from the model using the save_weights function.

4) *Control Application using predicted Hand gestures*

The trained model in JSON format and the model weights are loaded to predict the hand gestures. The PyAutoGUI which is used for Keyboard key integration with hand gestures and Streamlit which is used to create a user interface are also imported.

When the user clicks the start button on the web page, the web camera starts running. The user can perform the hand gestures in the Region Of Interest and the trained CNN model will predict the gesture. Each hand gesture is integrated with a Keyboard key using PyAutoGUI with help of conditional if-else statements to call predicted gestures. Each gesture is mapped with a keyboard key control and a label. The number of presses is assigned as 1, so every time a gesture is predicted the integrated control function is performed once. The user can exit the system by pressing the escape keyboard key. The video frame will display the gesture predicted and the action being performed whenever the user is using the system to control the media player.

The following conditional statements are used to perform the actions as shown in Table I.

Predicted Gesture	Action	Keyboard Keys
Palm	PLAY/PAUSE	Play/pause
Fist	MUTE/UNMUTE	Mute
Thumbs Up	VOLUME UP	Volume up
Thumbs Down	VOLUME DOWN	Volume down
Index Left	FORWARD	Prevtrack
Index Right	REWIND	Nexttrack
No Gesture	NO-ACTION	NIL

TABLE I. GESTURES AND THEIR RESPECTIVE ACTIONS

B. Architecture

The outline of any gesture recognition system generally involves the following three aspects:

1. Data acquisition and pre-processing
2. Data representation or feature extraction
3. Classification or decision-making
4. The proposed system architecture can be seen in Fig. 1.5. It includes the three main aspects required in a gesture recognition system.

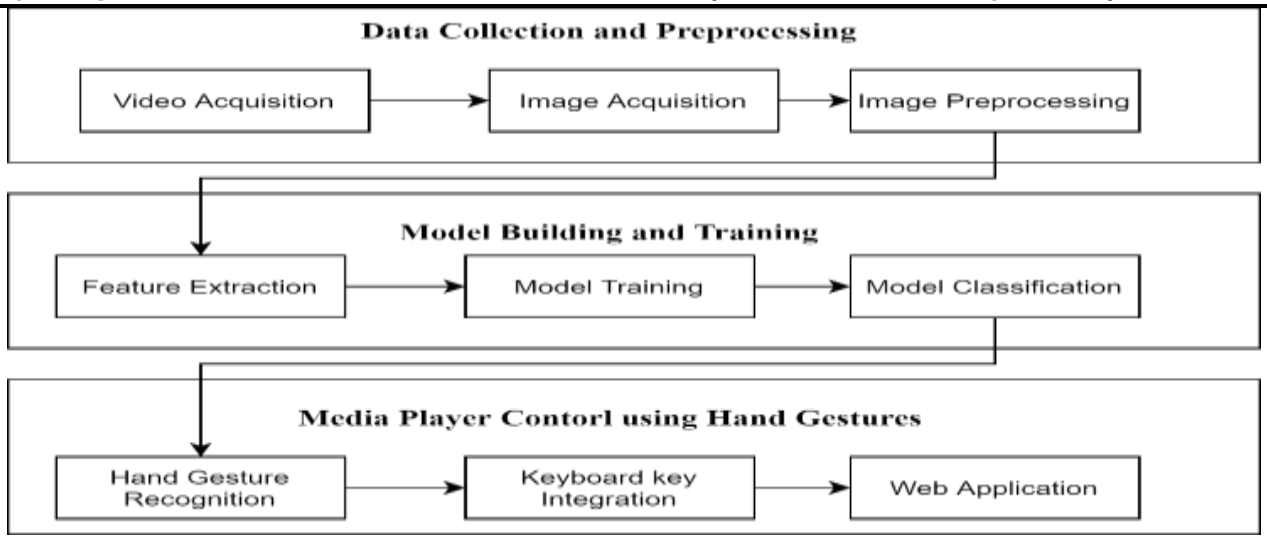

Fig. 1.5 Proposed architecture diagram

1) *Data Acquisition and Preprocessing Phase:*

The image data input is captured from a webcam, in the form of a video. The video is further broken into frames. These frames are converted into black and white images using OpenCV and then stored in specific.

2) Model Building and Feature Extraction:

A CNN model is built using Keras libraries. ImageDataGenerator class is used to pre-process the images in the directories to extract only the needed features. Model directories is trained using train dataset images.

3) Classification and Prediction:

Compilation of the model is performed and test accuracy is evaluated. Keras libraries are used to perform save and load of the model and to classify the gestures to a particular class.

4) Integrating the Keyboard controls:

Integrate every gesture with control functions using the Pyautogui library. The control of the prediction is executed.Deploy a web app with all project files using streamlit sharing.

IV. CONCLUSIONS

The Hand-gesture implementation involves significant usability challenges, including fast response time, high recognition accuracy, quick to learn, and user satisfaction.

The application of manipulation of objects through user hand gestures is being proposed and implemented in this paper that provides a suitable efficient and user-friendly human computer interface. With the aid of this application humans can interact with the virtual objects using hand gesture rather than any other physical input devices.

Users can handle multiple applications from distance using hand gestures without even touching it. This technique can be very helpful for deaf and dumb people because they can define the gesture according to their need.

The gesture would function as a direct command for operations like play or pause the video, volume up or volume down supporting the user's gestures onto the screen.

As a future prospect of this analysis we also tend/planning to investigate with the huge number of gestures with different people and motion type hand gestures will be developed. We are also going to generalize our system so that it can be useful for other different applications like power point presentation, games, windows picture manager etc.

V. ACKNOWLEDGMENT

We would like to express our gratitude to our principal Dr. Ravishankar M sir and our vice principal and HOD of CSE dept Dr. C Nandini mam for giving us an opportunity to showcase our project and research work and our beloved guide Dr. C Nandini mam for guiding us in innumerable ways towards the development and success of the project.

VI. REFERENCES

[1] Niranjani V, Keerthana R, Mohana Priya B, Nekalya K, Anantha Krishnan Padmanabhan. System application control based on Hand gesture using Deep learning. 7th International Conference on Advanced Computing & Communication Systems(ICACCS).20 Aug 2021.

[2] Gayathri Devi Nagalapuram, Roopashree S, Varshashree D, Dheeraj D, Donal Jovian Nazareth. Controlling media player with hand gestures using convolution neural network. IEEE Mysore Sub Section International Conference. Apr 2021.

[3] Parshwa P. Patil, Maithili J. Phatak, Suharsh S. Kale, Premjeet N. Patil, Pranav S. Harole, Prof. Dr. C. S. Shinde. Hand Gesture Recognition for Deaf and Dumb. International Research Journal of Engineering and Technology (IRJET).Volume:06, 11 Nov 2019.

[4] B.Kiruthiga, R.Narmatha Banu, K.M.Aparna, J.Deetchith Kaiserea, D.Selva Kumari. Gesture Control Smart System for Deaf and Dumb People. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering.Vol. 8, Issue 3, March 2019.

[5] R.Jagadish, R.Gayathri, R.Mohanapriya, R.Kalaivani, S.Keerthana. Hand Gesture Recognition System for Deaf and Dumb Persons. Indo-Iranian Journal of Scientific Research (IIJSR).Volume 2, Issue 1, Pages 139-146, January-March 2018.

[6] Renuka.N , Shyalu .S , Ramalakshmanan .B ,Selvakumar. S. Hand Gesture for Multiple Application. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056. Volume: 03 Issue: 02 | Feb-2016.

[7] Neha S.Rokade, Harsha R.Jadhav, Sabiha A.Pathan, Uma Annamalai.Controlling Multimedia Application Using Hand Gesture Recognition.International Research Journal of Engineering and Technology (IRJET). Vol 2,05-Aug-2015.

[8] N.Krishna Chaitanya, R.Janardhan Rao. Controlling of windows media player using hand recognition system. The International Journal of Engineering And Science(IJES). 01-Apr 2014.

[9] Zhuang H, Yang M, Cui Z, Zheng Q. A method for static hand gesture recognition based on non-negative matrix factorization and compressive sensing. IAENG International Journal of Computer Science. 2017 Mar 1;44(1):52-9.

[10] Yashas J, Shivakumar G. Hand Gesture Recognition: A Survey. In2019 International Conference on Applied Machine Learning (ICAML) 2019 May 25 (pp. 3-8). IEEE.

[11] Zhuang H, Yang M, Cui Z, Zheng Q. A method for static hand gesture recognition based on non-negative matrix factorization and compressive sensing. IAENG International Journal of Computer Science. 2017 Mar 1;44(1):52-9.

[12] Zhao L. Gesture Control Technology: An investigation on the potential use in Higher Education. University of Birmingham, IT Innovation Centre: Birmingham, UK. 2016 Mar.