Intelligent Media Player

Aakanksha Kapure[1], Prajakta Gole[2], Tejal Deshmukh[3]

Project Guide: **Mrs. Vidya Pujari**

Dept. of Information Technology, Vivekanand Education Society's Institute of Technology, Maharashtra, India

---***---

Abstract - *This paper aims to propose a desktop application that will able to recognize the user's face and hand, and according to that the video will be played or paused on a player. With the advancement in technology, now we can develop an application where the user just has to do eye or face movements or hand gestures and according to that player will act. We can develop this project by using a popularly adopted domain by researchers - Artificial Intelligence (AI). Artificial Intelligence gives birth to several applications where a system can be used to work as a human expert. This application can be developed using various algorithms such as Haar Cascade Classifier. Nowadays everything is online. This player will play a major role in all the platforms as it will play and pause depending upon the user actions.*

***Key Words*: artificial intelligence, desktop application, face movements, haar cascade classifier, hand gestures, human expert**

1. INTRODUCTION

Intelligent Media player is a player which is designed for time-saving as well as for the handicapped people who are not able to do the movement of the hands. This player will play an essential role as it includes eye, face, and hand gestures according to which player will be able to play and pause the video. Usually, when you are watching a video and someone calls you, you have to look somewhere else or go away from the screen for some time so you miss some part of the video. Then need to drag back the video from where you left it.

To overcome this problem, we will try to develop a look-based media player that pauses itself when the user is not looking at it. The player will play the video after the user looks at the screen again. For this, we need the camera or webcam on top of the computer. As long as the camera detects the user's face, eyes or hand gestures the video will be played. The player will pause as soon as the user's face, eye, or hand gestures is not completely recognized.

This paper is organized as follows: Section 2 explains the need and scope of the project. Section 3 has the literature survey. Methodology in Section 4 followed by the flow diagram in Section 5. Section 6 talks about the result analysis followed by conclusions drawn from it in section 7.

2. NEED AND SCOPE

Following is the need and scope of the project:

- It will be useful for people with certain disabilities who can play and pause the video by eyes and face.
- Due to the pandemic; from nursery to big industries everything became online.
- This player will play a major role in all the platforms as it will play and pause depending upon the user actions.
- It will also be useful for watching recorded lectures for students, workshops, and meetings for employees.

3. LITERATURE SURVEY:

3.1. MP-FEG: Media Player controlled by Facial Expressions and Gestures. [1]

This paper helped us to understand a new technique to interact with the computer in a non-tangible way. This Media Player system controller is designed for Facial Expressions and Gestures (MP-FEG). They have used Tangible devices like keyboard and mouse to give instructions to the computer for a long time. The next step in the development of human-computer interaction (HCI) is a non-tangible way of communication between humans and computers. This system is to be built using a perspective user interface in the type of HCI system. The main purpose is to find a non-tangible way to interact with the computer and for this, they have experimented to verify whether the facial expressions and the hand gestures can be used to give a command to the computer, specifically for controlling a media player system in the real-time situation. For facial expression, they extracted landmark points using an efficient discriminative deformable model. This discriminating model gives 49 landmark points on the face region. To handle different positions, rotations, and scales of the face in the image in different videos, the extracted facial landmark points are aligned to a reference shape using Procrustes analysis.

3.2. An Approach for face detection using Artificial Intelligence. [2]

This paper proposes a method to detect faces in a given image by the addition of a Gabor filter and Neural network. In the first phase, they used a Gabor filter which generates a feature set. Face and non-face templates are taken and processed with a Gabor filter. In the spatial (time) domain, the face images are present. The conversion of images into the frequency domain is processed through inverse fast Fourier transform. The subsequent frequency domain images are conjugated with a Gabor filter bank and a feature vector is generated.

The second phase involves a method where all the features are given as input to the neural network of two hidden layers with scaled conjugate training. Thus, this approach being deployed is a convolution of Gabor filter with frequency domain of training and test images provided a feature vector that was sourced to the neural network.

3.3 Rapid Object Detection using a Boosted Cascade of Simple Features. [3]

We learned about a machine learning method for visual object detection in this paper, which is capable of processing images quickly and achieving high detection rates. Three main contributions distinguish this work. The first is the introduction of a new image representation known as the "Integral linage," which allows our detector's features to be computed quickly.

The second is an AdaBoost-based learning algorithm that selects a small number of key visual features from a larger collection to produce extremely efficient classifiers.

The third contribution is a method for combining increasingly more complex classifiers in a "cascade" that allows background regions of the image to be rapidly discarded while more computation is spent on promising object-like regions. The cascade can be assumed of as an object-specific focus-of-attention technique that, unlike previous methods, offers statistical guarantees that discarded regions will not be re-examined.

The system achieves detection rates compared to the best previous systems in the face detection field. The detector, which is used in real-time applications, operates at 15 frames per second without using image differencing or skin colour detection.

3.4. Human face detection algorithm via Haar cascade classifier combined with three additional classifiers. [4]

From this paper, we learned about a new human face detection algorithm based on cascade classifiers using Haar-like features. Three supplementary weak classifiers are appended to the primitive Haar-like features-based cascaded classifiers.

First is a decision node based on human skin histogram matching.

The second weak classifier is based on eye detection and the third weak classifier is based on mouth detection. Eyes and mouth detections are also implemented with Haar-like feature-based cascade classifiers, hence; both of them have a sufficiently high detection rate, which satisfies conditions of weak classifiers.

Two supplementary classifiers based on eyes and mouth detections further remove those non-faces whose colors happen to be under the human skin color, but there are probably no eyes- and mouth-like objects in it.

Due to the availability of modules in OpenCV proposed human face detection system is simple to implement.

3.5. Real-time Vision-based Hand Gesture Recognition Using Haar-like Features. [5]

From this paper, we precisely understood a two-level approach to solving the problem of real-time vision-based hand gesture classification. Posture recognition with Haar-like features and the AdaBoost learning algorithm is implemented via a lower-level approach. Using this algorithm, real-time performance and high recognition accuracy can be obtained. The linguistic hand gesture recognition using a context-free grammar-based syntactic analysis is implemented at a higher level. Given an input gesture, based on the extracted postures, the composite gestures can be parsed and recognized with a set of primitives and production rules.

A parallel cascade structure is implemented to classify different hand postures, based on the cascade classifiers. From the experiment results, we conclude that this structure can achieve satisfactory real-time performance and also; very high classification accuracy above 90%. For high-level hand gesture recognition, they proposed context-free grammar to analyze the syntactic structure which is based on the detected postures.

3.6. Human Computer Interface Using Hand Gesture Recognition Based on Neural Network. [6]

From this paper, we learned about a hand gesture interface for controlling a media player using a neural network. This algorithm recognizes a set of four specific hand gestures, which are: Stop, Play, Reverse, and Forward.

This algorithm is based on four phases: Feature extraction, Image acquisition, Hand segmentation, and Classification. An image frame from the webcam will be captured, and then skin detection will be used to segment skin regions from background pixels. A new image is created which contains a hand boundary. For describing the hand gesture, hand shape features extraction is used. An artificial neural network is used as a gesture classifier. For training 120 gesture images have been used. The obtained average classification rate is 95%. This algorithm develops an alternative input device to control the media player, and also offers different gesture commands which can be useful in real-time applications.

4. METHODOLOGY

The algorithm which we have used is Haar-Cascade Classifier

Haar Cascade Classifier:

- Haar Cascades classifiers are an effective way for object detection. This method is proposed by Paul Viola and Michael Jones in their paper "Rapid object detection using a boosted cascade of a simple feature."

- Haar cascade is a method in which a lot of positive and negative are used to train the classifier.

- Positive images contain the actual objects which we want our classifier to detect.

- Negative images contain everything that we do not want to detect.

- The first step is to gather the Haar features. A Haar feature is a set of calculations performed on adjacent rectangular regions in a detection window at a particular location. The calculation includes adding up the pixel intensities in each section and figuring out the differences between the sums.

- The features can be difficult to spot in a large image. Integral images are useful in this situation. The calculation of these Haar features is quickened by using integral images. Rather than computing at each pixel, it divides the screen into sub-rectangles and generates array references for each of them. The Haar features are then computed using these.

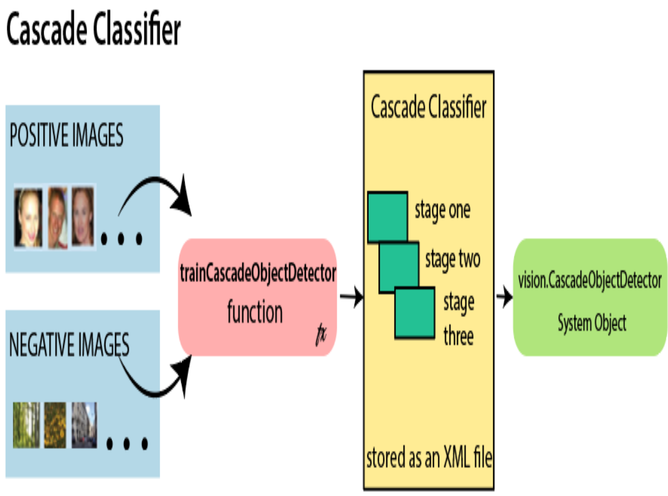

Figure 1: Haar Cascade Classifier

- Images features are treated as numerical information extracted from pictures that can distinguish one image from another.

- We apply every feature of the algorithm to the training images. At the start, every image is given equal weight. As result, it finds the best threshold which will categorize the faces to positive and negative.

- There may be errors and misclassifications. We select features with a minimum error rate, which means these are the features that best classifies the face and the non-face images.

5. FLOW DIAGRAM

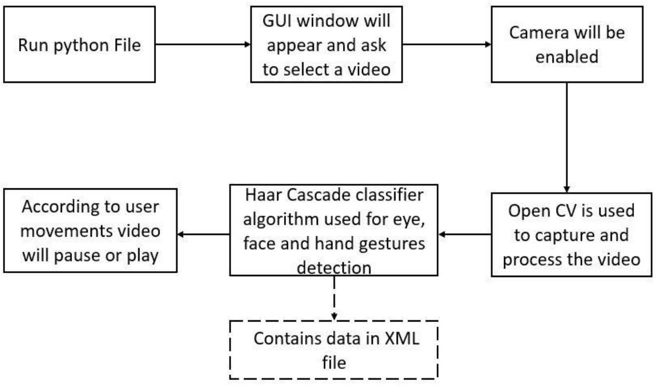

Figure 2: Flow Diagram

- First, we will run the python file(.py).
- After running the python file GUI window will appear and asks you to select the video from a local device.
- After selecting the video; the camera will enable and, the video will start playing. The camera will detect the eye, face and, hand movements while the video is playing.
- OpenCV is used for image processing also it is used to detect the camera. In our project, we used the open cv python library to capture video. OpenCV already contains many pre-trained classifiers for eyes.
- Face detection using Haar cascade classifier algorithm in a machine learning-based approach where a cascade function is trained with a set of input data. Data contains in XML files.
- When the user closes his eyes, the video will pause, and when the user opens his eyes video will resume. When the user looks in a different direction (i.e., the face is not fully visible) video will stop. When the user shows palm video will stop playing and when the user shows fist of hand video will resume.

6. RESULT ANALYSIS

1) After running the python file (Main.py). The GUI window will appear and asks you to select the video from a local device.

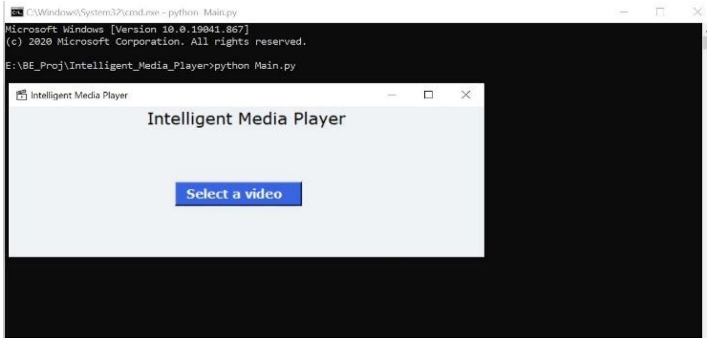

Figure 3: Window for selecting a video

2) Once we click on select video; video files on the computer will be shown:

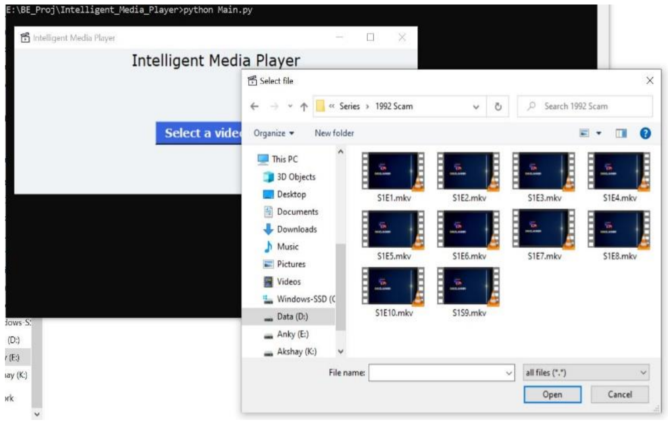

Figure 4: Multiple video file options

3) After selecting the video; the camera will enable and the video will start playing. The camera will detect the face and eyes as shown in the following image green square indicates eyes and red square indicates the face.

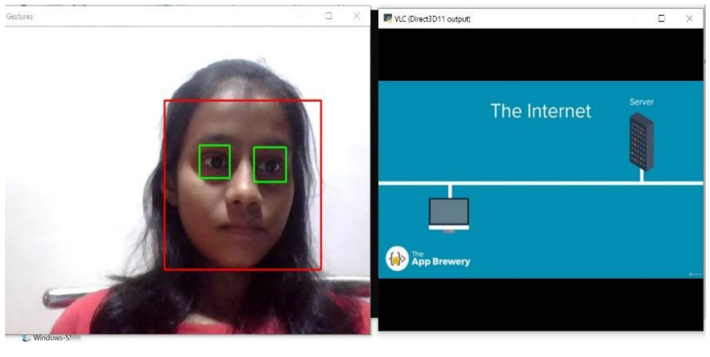

Figure 5: Testing Eye Detection Feature - A

4) As shown in the following output camera will only detect the face when eyes are closed; if the camera is not able to detect the user's eyes video will pause.

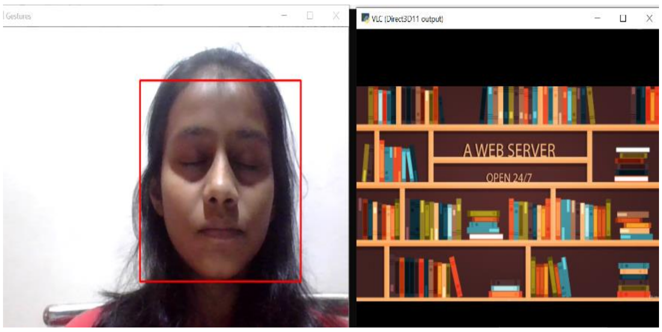

Figure 6: Testing Eye Detection Feature - B

5) As shown in the following output if the camera is not able to detect face, eyes and, hand gestures the video will stop playing.

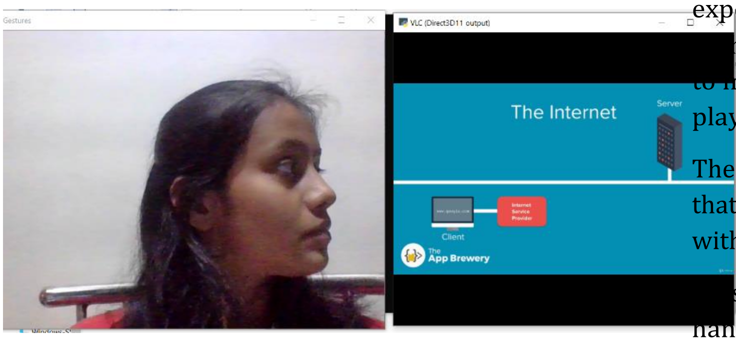

Figure 7: Testing Face Detection Feature

6) In the following output video is paused; using hand gestures if the camera detects the palm video will pause.

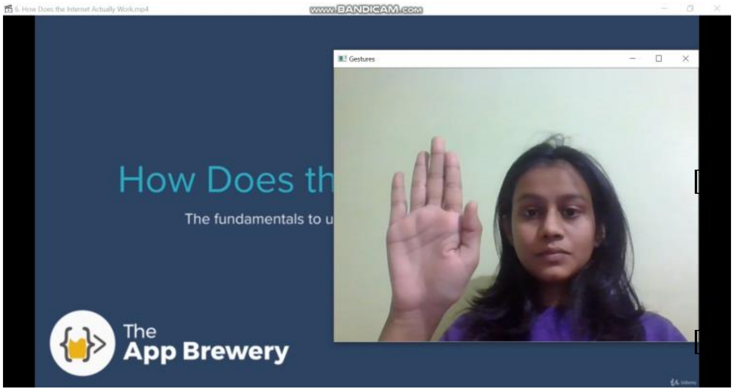

Figure 8: Testing Hand Gesture by showing palm

7) If the camera detects the fist of hand video will be played.

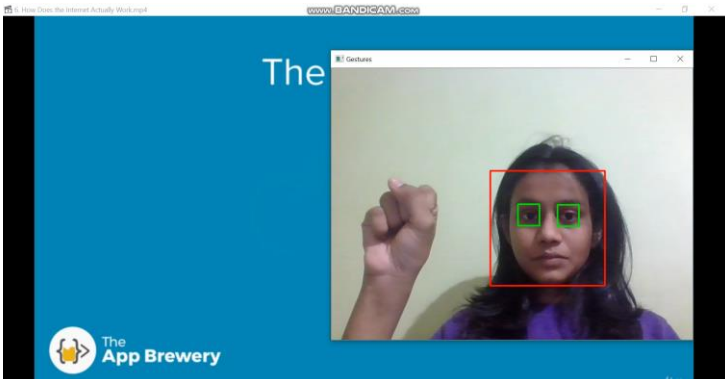

Figure 9: Testing Hand Gesture by showing Fist

7. CONCLUSIONS

In this project, we aim to help the user to get a better experience of media player. When the user is not looking he screen, we use hand gestures and face recognition to monitor the functionality of the media player such as playing and pausing the video itself.

The main investigation of this paper was exploring a system that will be able to recognize the face and hand gestures with a user-friendly GUI

s player will be very useful for paralyzed and handicapped people as it would detect the user's face and hand movements.

REFERENCES

[1] Swapna Agarwal, Saiyed Umer: "MP-FEG: Media Player controlled by Facial Expressions and Gestures" IEEE National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG) - 19 Dec. 2015

[2] Vandana Bhat, Dr. Jagdeesh D. Pujari: "An approach for face detection using artificial intelligence" 2014 International Conference on Contemporary Computing and Informatics (IC3I) - 29 Nov. 2014

[3] Paul Viola, Michael Jones: "Rapid-object-detection-using-a-boosted-cascade-of-simple-feature" Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 - 14 Dec. 2001

[4] Li Cuimei, Qi Zhiliang, Jia Nan, Wu Jianhua: "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers" 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI) - 22 Oct. 2017

[5] Qing Chen, Nicolas D. Georganas, Emil M. Petriu: "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features" 2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007 - 3 May. 2007

[6] Hamid A. Jalab, Herman.K. Omer: "Human Computer Interface Using Hand Gesture Recognition Based on Neural Network" 2015 5th National Symposium on Information Technology: Towards New Smart World (NSITNSW) - 19 Feb. 2015

[7] Himadri Nath Saha, Sayan Tapadar, Shinjini Ray, Suhrid Krishna Chatterjee, Sudipta Saha: "A Machine Learning Based Approach for Hand Gesture Recognition using Distinctive Feature Extraction" 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC) – 10 Jan. 2018

[8] Laura Liu and David Zhang: "Palm-Line Detection" IEEE International Conference on Image Processing 2005 – 14 Sept. 2005

[9] Mathias Kolsch, Matthew Turk: "Robust hand detection" Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings – 19 May. 2004

[10] Huda Abualola Hanin Al Ghothani, Abdulrahim Naser Eddin, Nawaf Almoosa Kin Poon: "Flexible Gesture Recognition Using Wearable Inertial Sensors" 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS) - 19 Oct. 2016