Microsoft Intern Engage 2022

PROJECT SYNOPSIS

Problem Statement Chosen : Face Recognition

Title of the Project

AVISHKAR – The Face Detection System

Submitted By

Utkarsha Avirat Sutar

Email ID : utkarshasutar19@gmail.com

Year : 2021-2022

Under the Guidance of

Mr. Amit Kumar Gupta

Microsoft, Hyderabad

Acknowledgement

I express our deep sense of gratitude to our respected guide Mr. Amit Kumar Gupta Sir for his valuable help and guidance. We are also truly thankful to them for encouragement they have given towards the planning of our Project till its execution.

I am also thankful of to Microsoft for providing such a beneficial opportunity to students. I also express my gratitude towards the Acehacker Team for conduction and scheduling of this programme.

At the last, I would like to express deep appreciation towards my mates at Microsoft Engage 2022 and my college mates and indebtedness to our parents for providing us the moral support and encouragement.

Abstract

The Proposed System of Avishkar – The Face Detection System is developed for providing an ease to its users. Multiple modules like Hand gestures Detection, Mask Detection, Emotion Detection, Counting Number of People in a frame and Object Detection are implemented in this prototype in order to make it a standalone application. Along with providing ease to its users, it also helps differently abled people to work efficiently and have a support at every stage of their life. This System is very User Friendly along with an interactive User Interface.

Keywords : System, Modules, Uses, User Friendly, Interactive Interface

Table of Contents

Introduction

Avishkar – The Face Detection System in which multiple modules are included in order to make one standalone application. It is a prototype based on Python that uses multiple features and packages from python in order to implement all of those features. This project contains a home page that redirects us to the modules implemented in it. The modules that are implemented in this project are : Hand Gestures Detection, Mask Detection, Counting Number of People in a frame, Emotion Detection and Object Detection.

Additionally I provides an ease to its users through its user interface. In addition, it also helps differently abled people to carry on their daily chores at an ease. This application is interactive and also speaks to its user for telling what's next and what is its use. Its main goal is to perform operations and assigned task with an ease that provides an efficiency to the user who is handling the application. This project is developed in Python and uses various libraries in order to operate the functionalities.

Literature Survey

1. Shalini Gupta ETL [2016] presents a novel recurrent 3D convolutional neural network classifier for dynamic gesture recognition. It supports online gesture classification with zero or negative lag, effective modality fusion, and training with weakly segmented videos. These improvements over the state-of-the-art are demonstrated on a new dataset of dynamic hand gestures and other benchmarks. [1]

2. Preeti Nagrath ETL [2021] presents in proposed face mask detection model named SSDMNV2, both the training and development of the image dataset, which was divided into categories of people having masks and people not having masks have been done successfully. The technique of OpenCV deep neural networks used in this model generated fruitful results. Classification of images was done accurately using the MobilenetV2 image classifier, which is one of the uniqueness of the proposed approach. [2]

3. V. Aparna ETL [2015] People counting has a wide range of applications in the context of pervasive systems. These applications range from efficient allocation of resources in smart buildings to handling emergency situations. There exist several vision based algorithms for people counting. Each algorithm performs differently in terms of efficiency, flexibility and accuracy for different indoor scenarios. Hence, evaluating these algorithms with respect to different application scenarios, environment conditions and camera orientations will provide a better choice for actual deployment. For this purpose, in our paper the most commonly implemented Frame Differencing, Circular Hough Transform and Histogram of Oriented Gradient based methods are evaluated with respect to different factors like camera orientation, lighting, occlusion etc. The performance of these algorithms under different scenarios demonstrates the need for more accurate and faster people counting algorithms. [3]

4. Shiv Naresh Shivhare ETL [2012] Emotion can be expressed in many ways that can be seen such as facial expression and gestures, speech and by written text. Emotion Detection in text documents is essentially a content - based classification problem involving concepts from the domains of Natural Language Processing as well as Machine Learning. In this paper emotion recognition based on textual data and the techniques used in emotion detection are discussed. [4]

5. Zhengxia Zou ETL [2019] Object detection, as of one the most fundamental and challenging problems in computer vision, has received great attention in recent years. If we think of today's object detection as a technical aesthetics under the power of deep learning, then turning back the clock 20 years we would witness the wisdom of cold weapon era. This paper also reviews some important detection applications, such as pedestrian detection, face detection, text detection, etc, and makes an in-deep analysis of their challenges as well as technical improvements in recent years. [5]

Problem Definition, Scope and Objectives

3.1 Problem Definition

"To develop easier and simpler Face Detection System which can help the user to handle the daily tasks and reduce the physical efforts and recognize human activities by accessing different modules in an Interactive way."

When it comes to people who are differently abled or illiterate people who find it difficult to read and understand, special feature is applied and implemented here in this project that reads out the content for its users.

3.2 Scope

This Proposed System is a windows base application which has capability to work with or without Internet Connectivity in desktop computer. It is a Face Recognition Technology, which uses the modules defined in it to perform various tasks and provide results to the end user.

There are five modules implemented in this prototype, they are :

Hand Gestures Detection : This model tells and specifies the user about what hand gesture is being done by any user, it may be like : live long, stop, okay, thumbs up, thumbs down, rock, fist, peace, smile.

Emotion Detection : This model determines exact emotions of a person whether he/she is neutral, happy, sad, angry etc.

Counting Number of People in a Frame : This system will return the count of people in a frame. It will detect all the people, count it and tell the result to the end user.

Mask Detection : In the situation of COVID-19, it is very essential to wear a mask. Although humans can't detect the ones who are not wearing a mask from a huge mob. So, in order to avoid it, Face Mask detection is used so that it can be detected efficiently.

Object Detection : This module will help us to detect the objects that appear into the frame like human, electronic gadgets, bottle etc. It will help us also so insecurity purpose like if we detect an intruder having knife in its hand it can tell the user to be alert.

3.3 Objectives

Following are the objective of proposed system. These objectives are going to achieved by developing this system :

1. To provide User Friendly Environment.
2. To provide ease and convenience to differently abled people and illiterate people in order to understand the application through sound generated through the application.

Details of design, working and processes
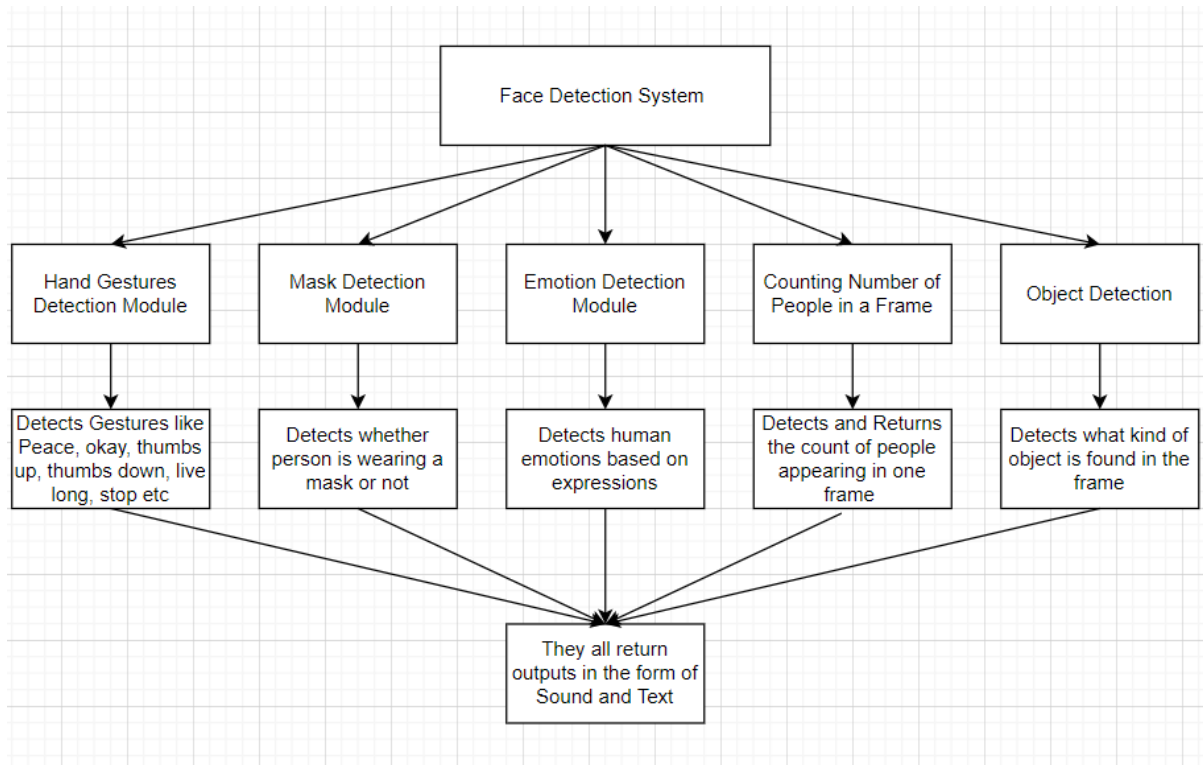
4.1 Design of Entire System

Fig. 4.1 Design of Face Detection System

There are five modules implemented in this prototype, they are :

Hand Gestures Detection : This model tells and specifies the user about what hand gesture is being done by any user, it may be like : live long, stop, okay, thumbs up, thumbs down, rock, fist, peace, smile.

Emotion Detection : This model determines exact emotions of a person whether he/she is neutral, happy, sad, angry etc.

Counting Number of People in a Frame : This system will return the count of people in a frame. It will detect all the people, count it and tell the result to the end user.

Mask Detection : In the situation of COVID-19, it is very essential to wear a mask. Although humans can't detect the ones who are not wearing a mask from a huge mob. So, in order to avoid it, Face Mask detection is used so that it can be detected efficiently.

Object Detection : This module will help us to detect the objects that appear into the frame like human, electronic gadgets, bottle etc. It will help us also so in security purpose like if we detect an intruder having knife in its hand it can tell the user to be alert.

4.2 System Requirements

1. Software Requirements
   a. Python 3.6 or above
   b. VS Code or any similar Text Editor Tool
   c. Python Packages
      - numPy
      - keras
      - tensorflow
      - PILLOW
      - tkinter
      - imutils
      - cv2
      - cmake
      - pyttsx3
      - mediapipe
      - argparse
2. Hardware Requirements
   a. PC with a Windows operating system (Windows 7, 8, 8.1 or 10).
   b. Minimum 4 GB of RAM. (Preferred 8 GB)
   c. 200 Mb of free space on hard drive.
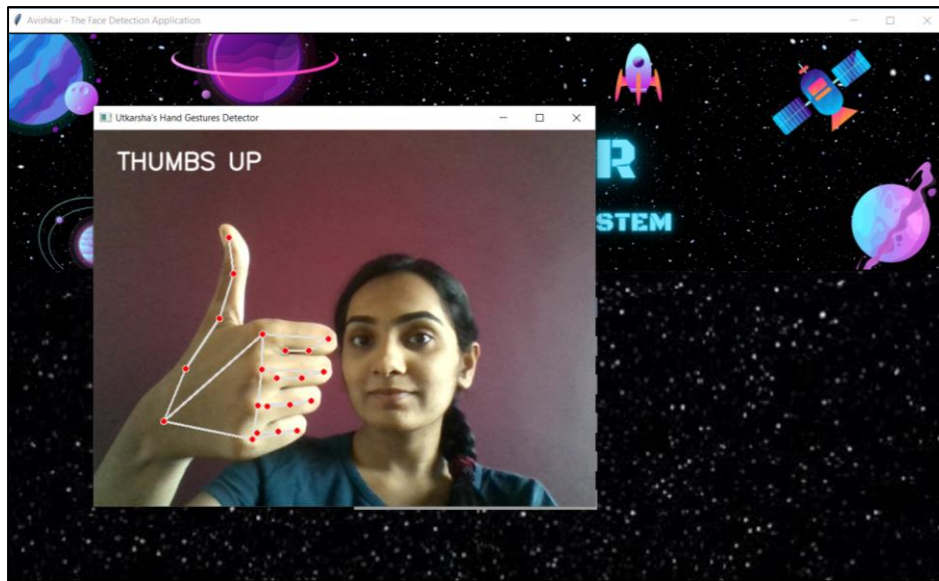   d. External USB microphone or headset.
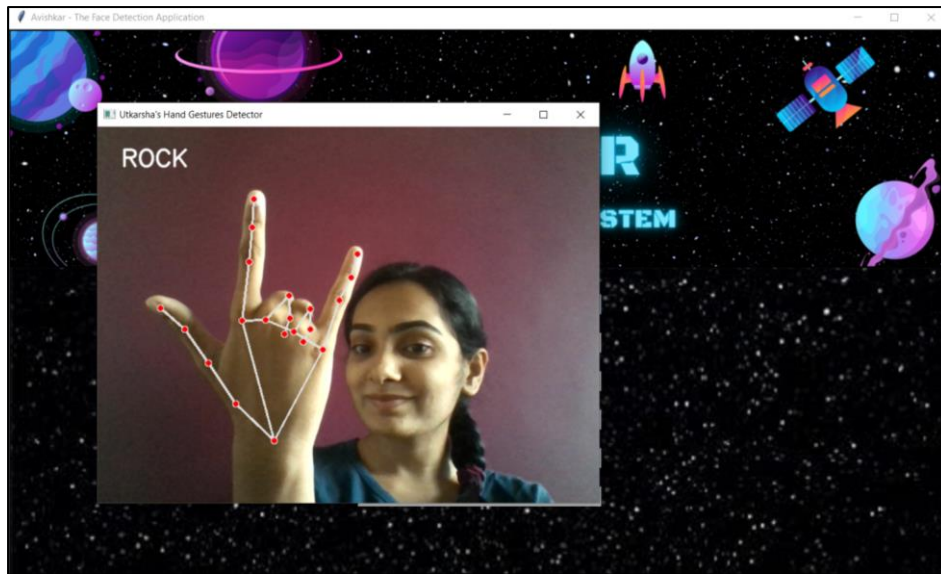
Results and Application
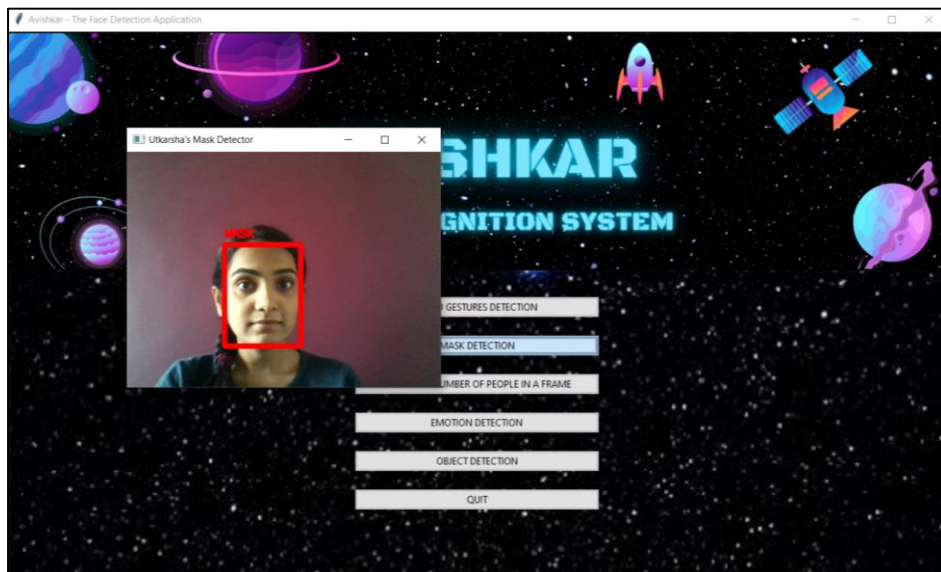
1. Homescreen / Main Page of the Application
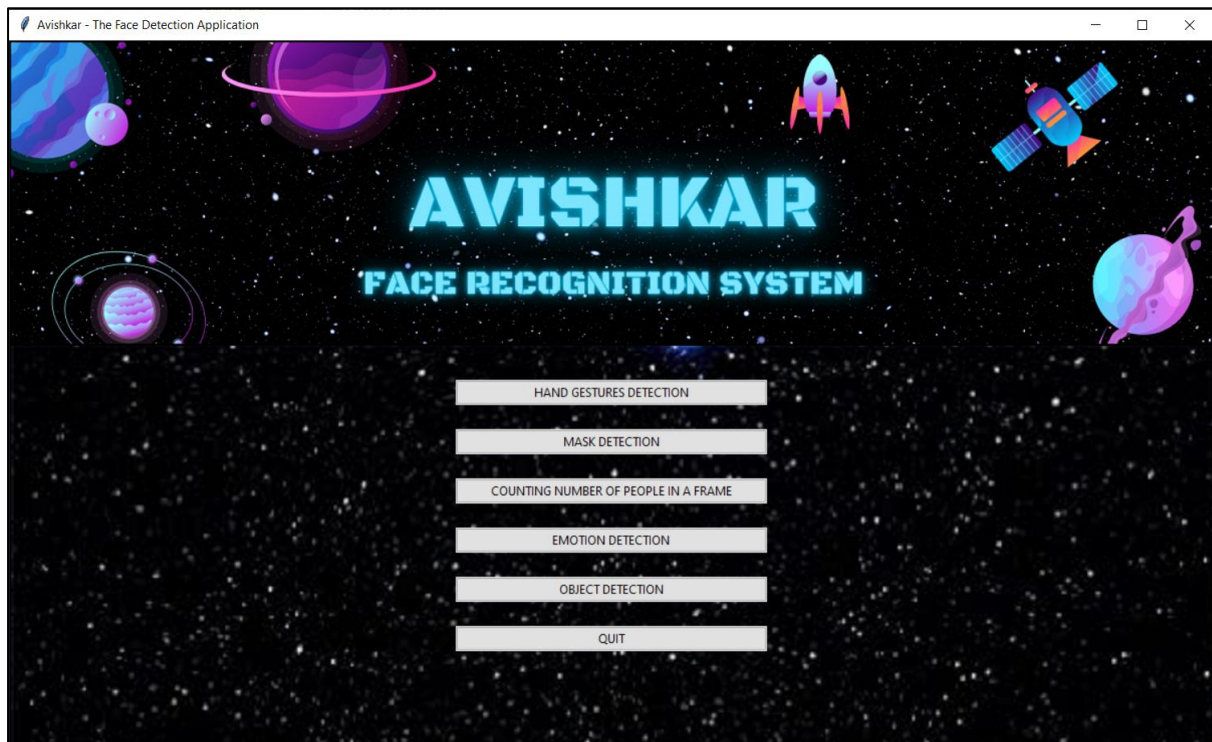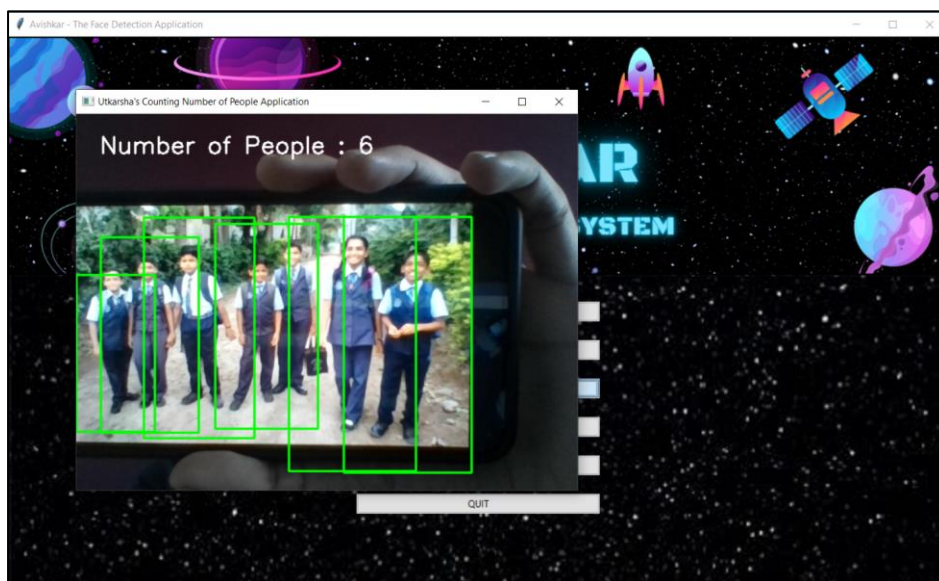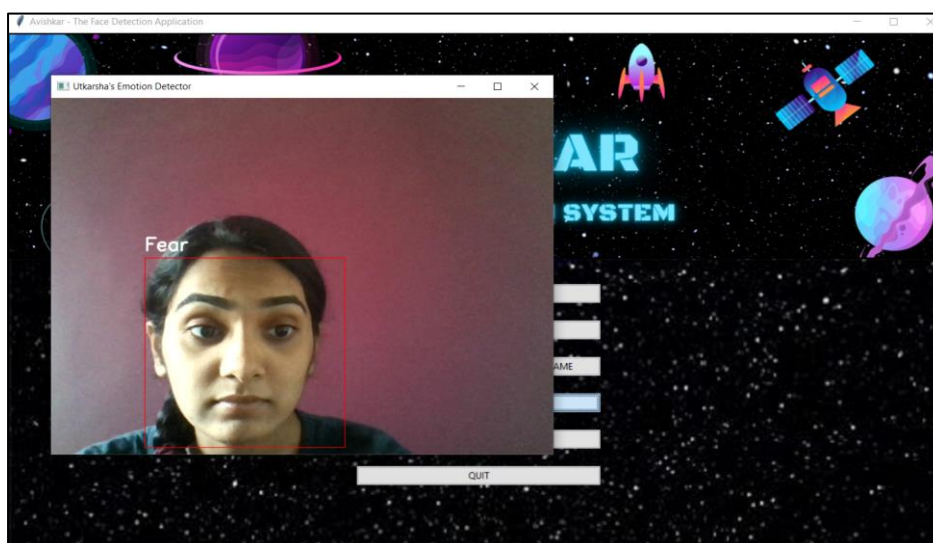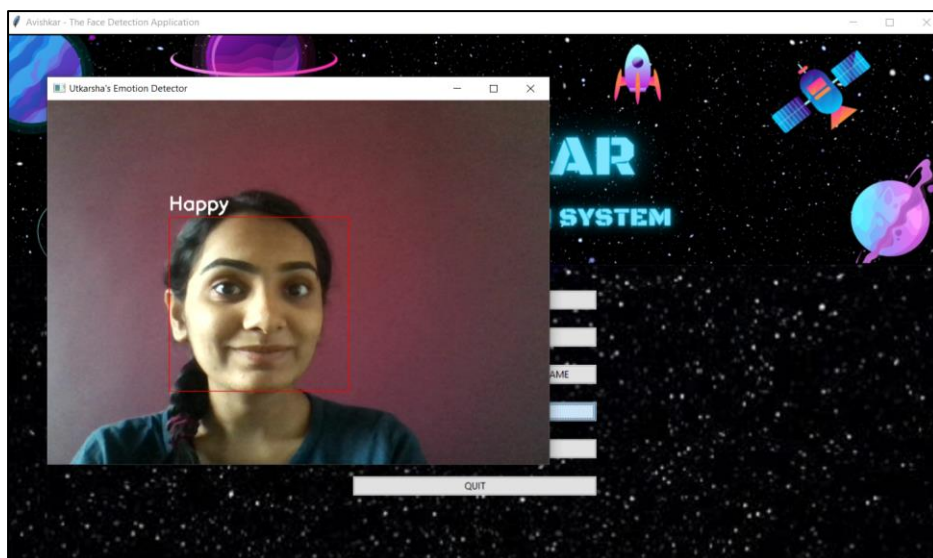

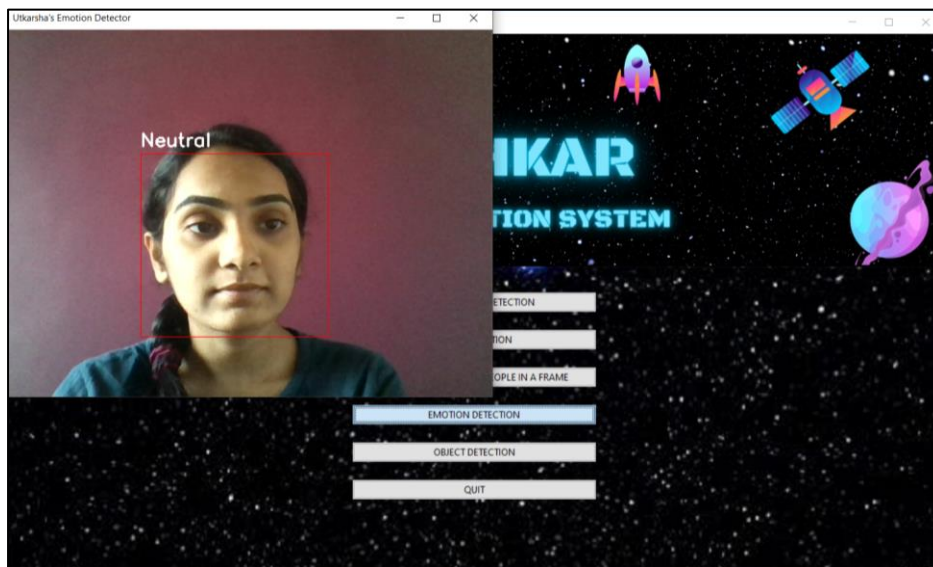2. Hand Gestures Detection Module
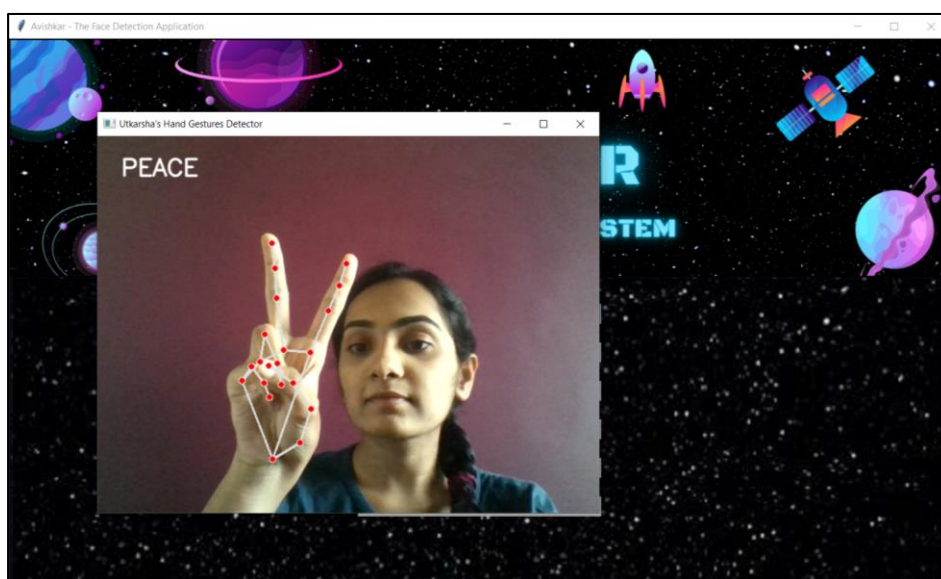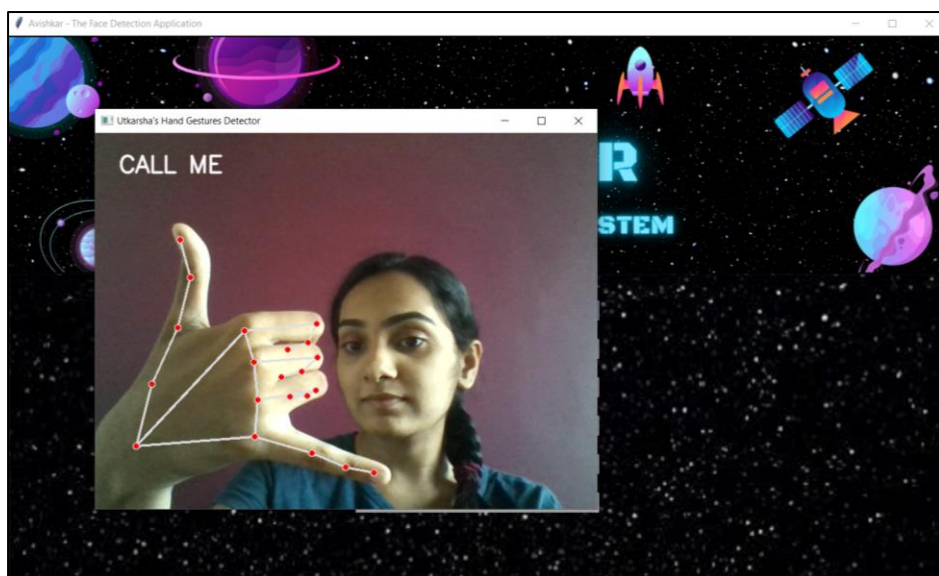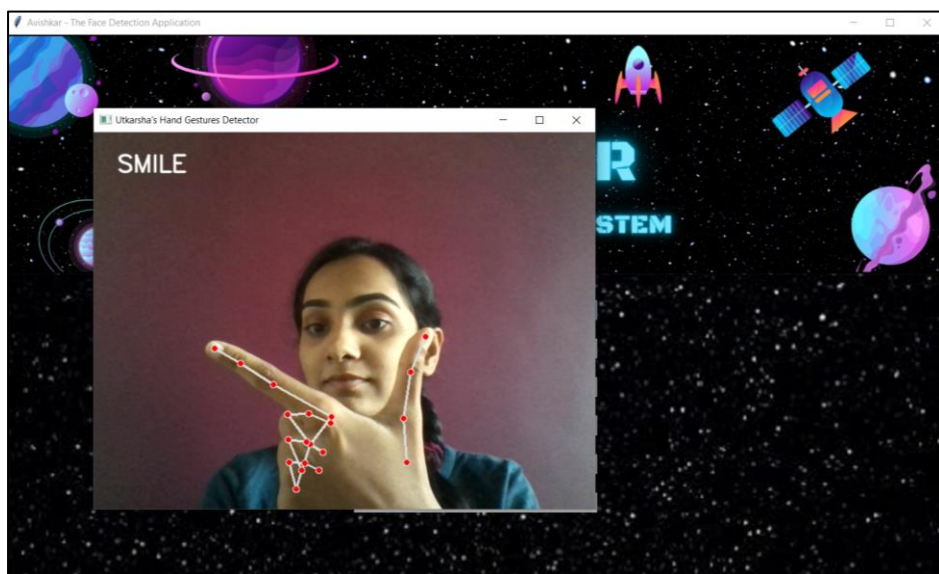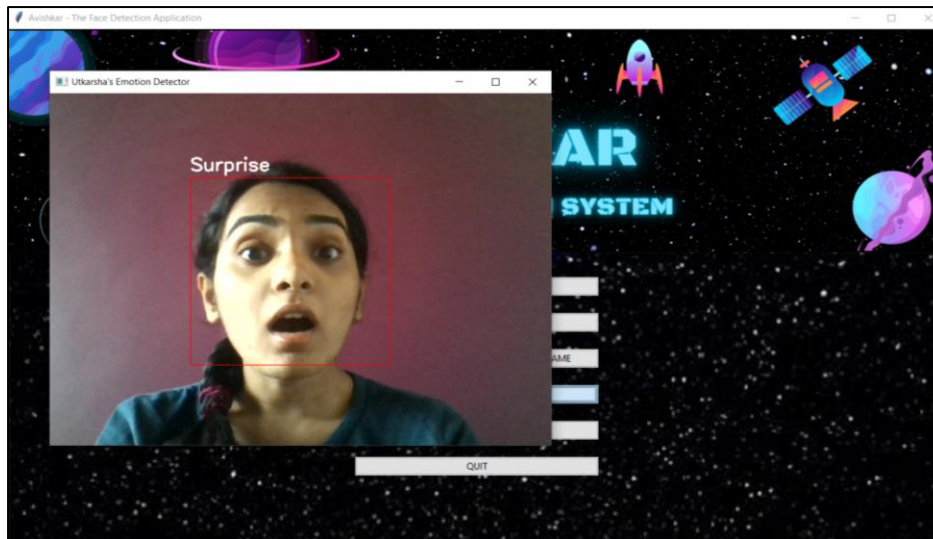
3. Mask Detection Module
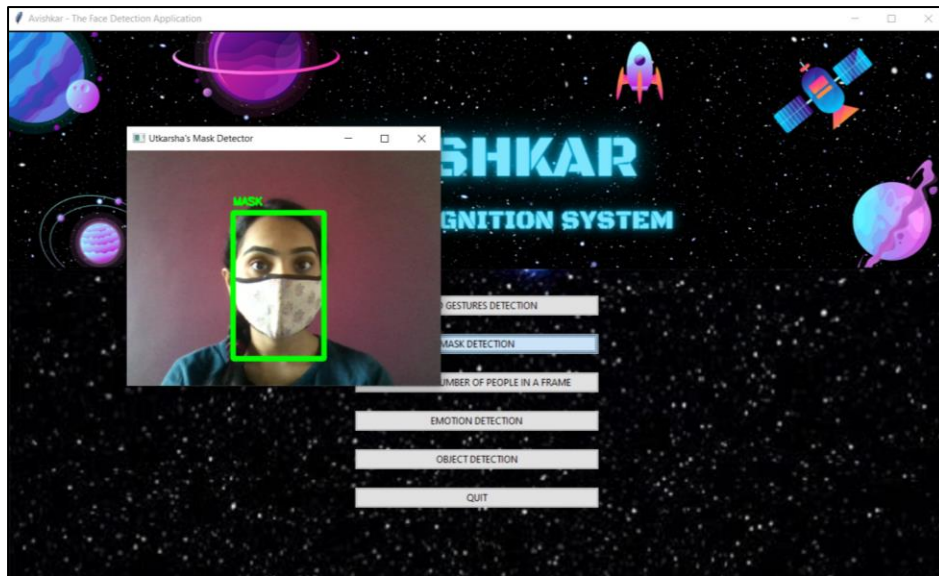
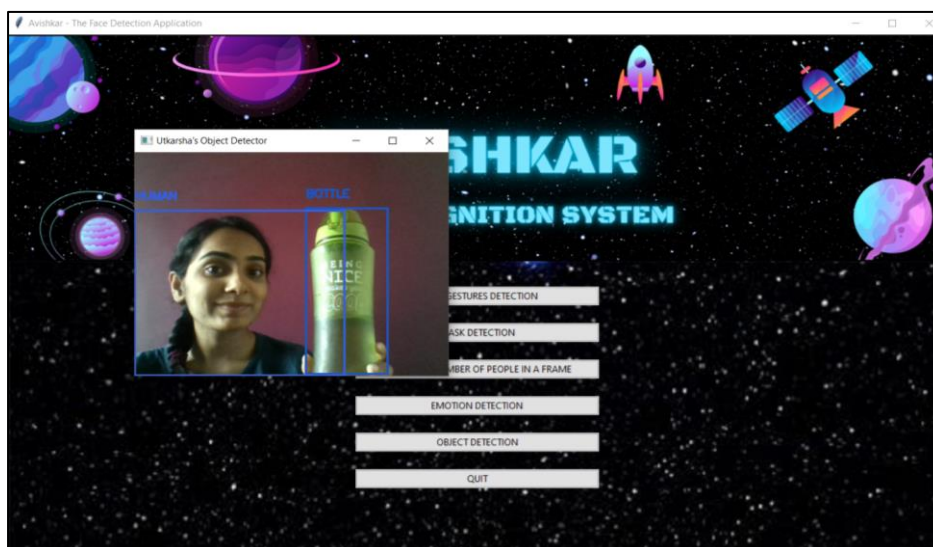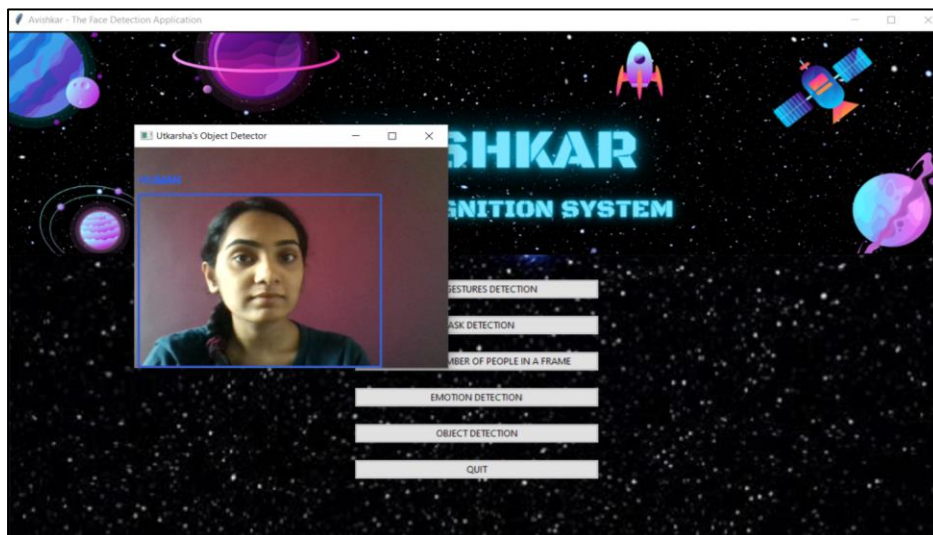4. Counting Number of People in a Frame

This system tries to return the result as an Approximate Answer

5. Emotion Detection Module

6. Object Detection Module

Applications

1. Helps Differently Abled people to carry on their daily work with almost no difficulties.
2. Helps Blind People by speaking the content.
3. Help Illiterate people to understand and interact with the application.
4. Application of Hand Gestures Detection
   - To understand the meaning of what a person is trying to convey.
5. Application of Mask Detection
   - To stop to spread of COVID-19 and strictly keep a track on people whether they are wearing a mask or not.
6. Application of Counting Number of People
   - For Online Examinations, if another person is detected in a frame, it may generate a warning.
7. Application of Emotion Detection
   - People who are not good at their Eyesight will also be able to know the feelings or expression of another person in front of them through sound.
8. Application of Object Detection
   - In security purpose, like if we detect an intruder having knife in its hand it can tell the user to be alert.

Conclusions and Future Scope

6.1 Conclusion

This prototype is developed for providing an ease of use to its users and also help to perform specific tasks. One of its applications like Object detection can be used for security purpose. Another one like the Counting number of people in the frame is used in online examinations. In such a way this application proves to be useful in various scenario.

6.2 Future Scope

1. User Authentication with the help of Face Recognition can be done.
2. Integration of some more features relate to face detection can be done.
3. Generating an Alarm on finding more people in a frame and intruder with a weapon can be done.

References and Bibliography

1. [Online Detection and Classification of Dynamic Hand Gestures With Recurrent 3D Convolutional Neural Network (thecvf.com)](#)

2. [SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2 | Elsevier Enhanced Reader](#)

3. [A Comparative Study of Vision Based Human Detection Techniques in People Counting Applications - ScienceDirect](#)

4. [[1205.4944] Emotion Detection from Text (arxiv.org)](#)

5. [[1905.05055] Object Detection in 20 Years: A Survey (arxiv.org)](#)