

BRIEFING

The project **Voice Command Based Object Detection** is made to detect object in an image or in real-time using one's webcam. This program works with both on the basis of voice assistance and with keyboard and cursor.

While running the program in first page user will receive 3 buttons.

The **first Button** having a shape of mic, “🎤” on left side of screen is to *turn on voice-based command*; where user can speak and give his command and even user will also receive feedback voice as whether your voice was clear and command was present.

The **second Button** is named as *Live object Detection* which gets access to webcam and start recognizing the object found in webcam, recognized objects are surrounded by a boundary of random color with name of object on the top of boundary and speaking the name of object.

The **third Button** is named as *Image Based Object Detection* which opens a new window so that *user will give path of the image file if wrong path is found user will receive an error message*. Later on, objects in this image are recognized and are named accordingly while speaking the name of objects captured in image. User must ensure that he should delete the file if not required or save in another folder if required.

While using this program one must ensure good internet connection so that voice assistance to work accordingly. Whenever an object is found we receive a voice message as the name of object detected.

This project is developed so that few objects among a group of objects can be classified. This can even act in creating a motion sensor as by changing a few lines of code and saving still image's

matrix and comparing it with real time image.

Technology Used

Python programming language is used in making of the project and several more modules were used in making of the project such as:

Tkinter Module:

```
from tkinter import *
```

This is used to create a GUI based window.

PIL Module:

```
from PIL import Image, ImageTk
```

This is used to display Image in GUI window.

CV2:

```
import cv2
```

This is Computer Vision Library which can be used to read image or video.

Numpy:

```
import numpy as np
```

This Library comes pre-installed with CV2 library this is generally for matrix calculation.

OS:

```
import os
```

This Library is imported to use run system command such as opening image, playing audio.

Speech Recognition:

```
import speech_recognition as sr
```

This Library is imported to recognize voice commands as it is basically speech-to-text conversion Library.

Pytsx3:

```
import pytsx3
```

This is a text-to-speech conversion library in python. Unlike other alternative Libraries, this library can work offline and is compatible with both Python 2 and Python 3.

Field of Project

This project can be used in places where high security is required or trespassing restricted areas. As if it detects some object this will alert you with voice assistance and even, we can use a simple few line of code to capture image and send it through E-mail to the deserving.

ADITYA PATHAK

LITERATURE SURVEY

Many other authors have created programs of object detection but this is different as it is voice assistance and can be implemented in real-life for house, office security etc. We receive an instant alert on object detection which will be beneficial for rapid action during some incident.

Already Running Projects on the basis of object detection

➤ OPTICAL CHARACTER RECOGNITION

This is the most widely used application of object detection, we make computer to recognize text on the basis of data provided, whether it's handwritten or typed.

This is used in Phones, OCR, Scanners etc.

➤ TRACKING OBJECTS

Object tracking has a variety of uses, some of which are surveillance and security, traffic monitoring, video communication, robot vision and animation.

Used in Motion Sensing System.

➤ SELF DRIVING CARS

Self-driving Cars requires object detection and we would essentially train the car to detect known set of objects such as cars, pedestrians, traffic lights, road signs, bicycles, motorcycles etc.

Used by Tesla, Google etc.

➤ FACE DETECTION AND FACE RECOGNITION

Face detection and Face Recognition is widely used in computer vision task. We noticed how Facebook detects our face when you upload a photo This is a simple application of object detection that we see in our daily life.

Used in Phones, Laptop, Certain Door locks.

➤ IDENTITY VERIFICATION THROUGH IRIS CODE

Iris recognition is one of the most accurate identity verification systems. Identity verification and identification is becoming increasingly popular.

Used in Samsung phones, restricted buildings etc.

➤ DIGITAL WATERMARKING

. It is typically used to identify ownership of the copyright of such signal. Digital watermarks may be used to verify the authenticity or integrity of the carrier signal or to show the identity of its owners. It is prominently used for tracing copyright infringements and for banknote authentication.

Used to avoid Copyright infringements.

➤ PEDESTRIAN DETECTION

Pedestrian detection is an essential and significant task in any intelligent video surveillance system, as it provides the fundamental information for semantic understanding of the video footages. It has an obvious extension to automotive applications due to the potential for improving safety systems.

Used by various Automobile Companies such as GM, Ford, Nissan etc.

METHODOLOGY

While making of project first thing is to gain access to camera, system files, sound etc. After that we need to perform a scan on the screen in multiple frames so that we get a location to search object for.

These are the main steps involved in image recognition.

- **Image Classification:** Predict the type or class of an object in an image.
 - *Input:* An image with a single object, such as a photograph.
 - *Output:* A class label (e.g., one or more integers that are mapped to class labels).
- **Object Localization:** Locate the presence of objects in an image and indicate their location with a bounding box.
 - *Input:* An image with one or more objects, such as a photograph.
 - *Output:* One or more bounding boxes (e.g., defined by a point, width, and height).
- **Object Detection:** Locate the presence of objects with a bounding box and types or classes of the located objects in an image.
 - *Input:* An image with one or more objects, such as a photograph.
 - *Output:* One or more bounding boxes (e.g., defined by a point, width, and height), and a class label for each bounding box.

When this is done, we need to initialize text-to-speech and speech-to-text so that it can understand user's command and act according to the command, after that we need to receive output from computer, this is done in both image format and voice format.

FACILITIES REQUIRED

- Laptop or Tablet or Smartphone
- Internet Connection
- Google Service
- A camera

ADITYA PATHAK