EXPERT SYSTEM FOR VISUALLY CHALLENGED USING AI

By

Adarsh verma(2147102),
Amber Ujjwal Linda(2147104),
Ankur Sharma(217105)

Project Guide
Dr. THIRUNAVUKKARASU V

A Project report submitted in partial fulfillment of the requirements for the award of degree of Master
of Computer Applications of CHRIST (Deemed to be University)
Dec- 2022

CERTIFICATE

*This is to certify that the report titled "**EXPERT SYSTEM FOR VISUALLY CHALLENGED USING AI**" is a bonafide record of work done by **Ankur Sharma (2147105), Adarsh Verma (2147102), Amber Ujjwal Linda(2147104)** of CHRIST (Deemed to be University), Bengaluru, in partial fulfillment of the requirements of V Trimester MCA during the year 2022.*

**Head of the
Department**

Faculty In charge

Valued-by:

Faculty In-charge:

Dr.THIRUNAVUKKARASU

Register Number(s)   :

2147105
2147104
2147102

CHRIST (Deemed to be University)

Examination Centre   :

Date   :

18 November 2022

ACKNOWLEDGMENTS

LIST OF FIGURES	
3.1 Solution Architecture diagram	4
3.2 Flow diagram	7
3.3 Algorithm and Data Structure	8
3.4 Entity Relationship diagram	9
3.5 Interface Design	10
3.6 Class Diagram	10

TABLE OF CONTENTS

ABSTRACT

This application helps people that are visually challenged to carry on the tasks for which they are dependent on someone else to be done. This application covers the processes such as text-to-speech conversion for office documents, library books, and any other printed text document and campus tours for the students' using audio output. This system can also be used for identifying any teacher based on the Face recognition concept and voicing out their details to the student. This complete system will majorly focus on the visually impaired students and help the students who are visiting the college for the first time. The development of this application aims to replace manual involvement with audio, and students do not need to rely on another person to complete the process.

1. <u>INTRODUCTION</u>

<u>1.1. PROBLEM DESCRIPTION</u>

Cell phones are an essential part of cutting-edge lifestyles. Proposed system is an android application that helps using voice commands. The utility is evolved for visually impaired human beings. After unlocking the mobile, the application can be launched with voice commands. The structure accept voice commands and volume button actions to perform operations according to them. For acting the further task, it first translates the voice into textual content, producing output in the form of voice. It performs simple features such as a campus tour of the college. We will use the concept of image processing to train our model to detect the different places inside the campus. This will help the visually impaired students to know about the campus, and by using this module, the visually impaired students will be independent enough to explore the campus on its own. In the next module, we have a concept of Teacher recognition which uses the concept of face detection, which uses image recognition and computer vision. The next module is Text to Speech recognition using which the student can scan any printed document and convert it to speech, and this module will support the students for documentation in the office and the library. The last module is about how the user will navigate through the landing page.

<u>1.2. EXISTING SYSTEM</u>

These are the apps which are made for visually challenged people namely Voice Over, Talk back, Be my eyes, Seeing AI, Aira, Lookout. These are the applications which use the concept of artificial intelligence and machine learning, they have most of the functions which are dependent on the mobile camera and the speaker. These components collect most of the information for processing the data. Some of the examples are text to speech, face recognition, reading the text out loud, detecting the object in front of them, scene descriptor and also scans the barcode of an item from a supermarket and gives the detailed information about the same. These functionalities use the concept of Natural language processing, sentiment analysis, computer vision and many more concepts.

1.3. PROPOSED SYSTEM

This system is going to be implemented as an android mobile application. It will consist of four functionalities that will be useful for the users who are visiting Christ university. These four functionalities are namely: a landing page, this page will direct the users to the other three functionality as the application is for users who are visually impaired it will ask for any input from the users side to select one of the functionality using a button or a gesture. The second functionality will consist of text to speech option for this the application will ask for permissions from the user which will ask to use the camera and microphone. This functionality will read any printed document and output the data in a audio format which was collected from camera. The next functionality is Teacher recognition, this is similar to face recognition in this the dataset few teachers will be trained from different angles and will be stored so that if a teacher from the collected dataset is in front of the camera then the application will respond with the teacher's name. The last functionality is the campus tour which will tell the user about the current place he is present.

2. SYSTEM ANALYSIS AND REQUIREMENTS

2.1. PROBLEM DEFINITION

Visually disbled people have a tough time finding good reading materials in accessible formats. Also the office forms are not easy to fill without anyone else's help. Other challenge that they face is recognising their teachers and they had to ask most of the time when they encounter them. Another challenge for the visually disbled person, especially the one with the complete loss of vision, is to navigate around places. Obviously, blind people roam easily around their house without any help because they know the position of everything in the house. But, unfortunately, this is not done in most of the places. This creates a big problem for blind people who might want to visit the place.

2.2. REQUIREMENT SPECIFICATIONS

2.2.1. FUNCTIONAL REQUIREMENTS

S.no	Functional Requirement
FR1	User should be able to Open the Application
FR2	User should be able to navigate through the three modules
FR3	User should be able to select the functionality
FR4	User should be able to hear the audio message regarding the choices
FR5	User should be able to use all the functionality
FR6	User should be able to go back to the landing page
FR7	User should be able to close the app even when the functionality is running

2.2.2 SYSTEM REQUIREMENTS

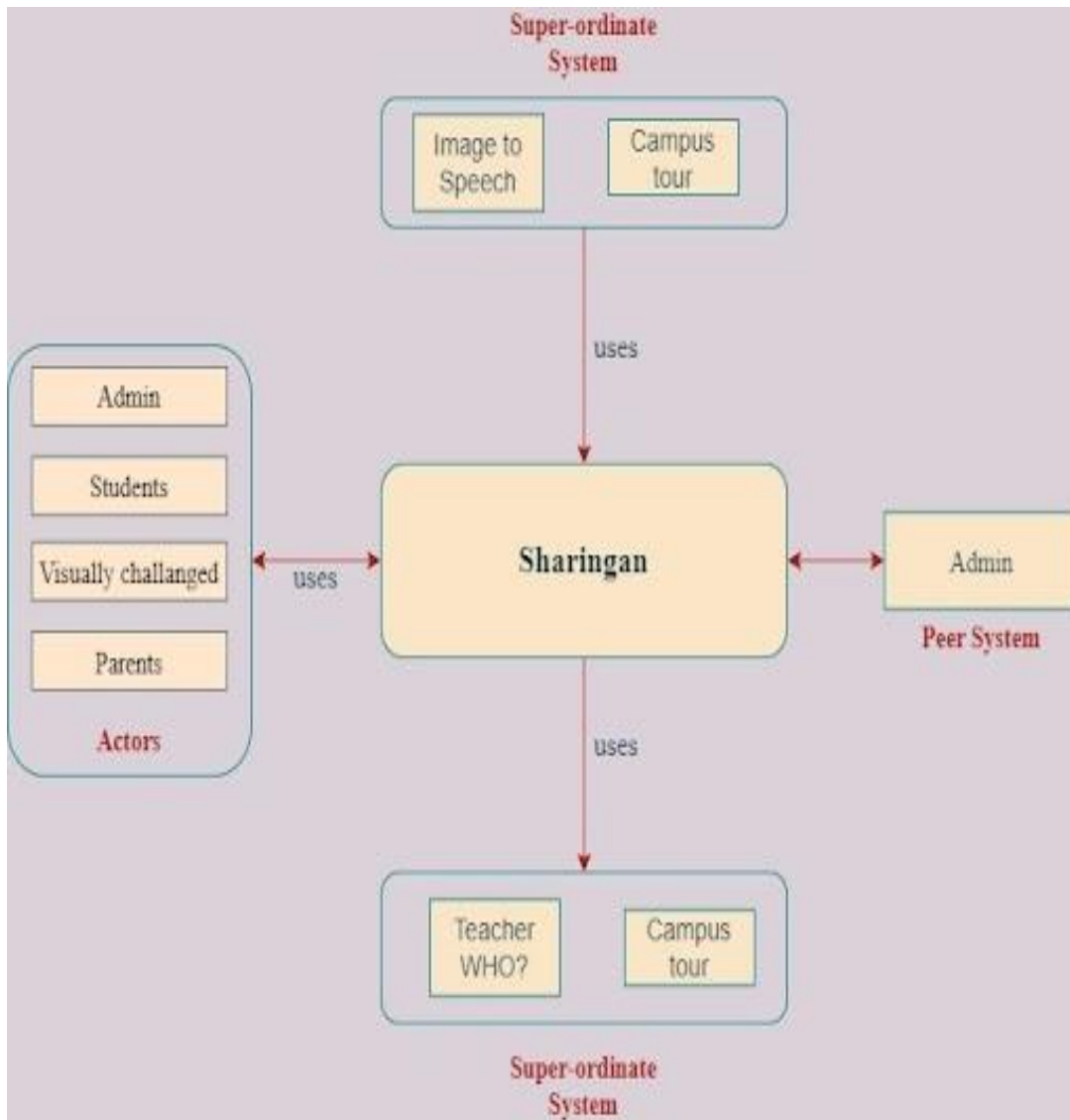**Hardware Requirements :**

- **Processor:** Intel Core i5 8th Gen
- **RAM:** 16 GB
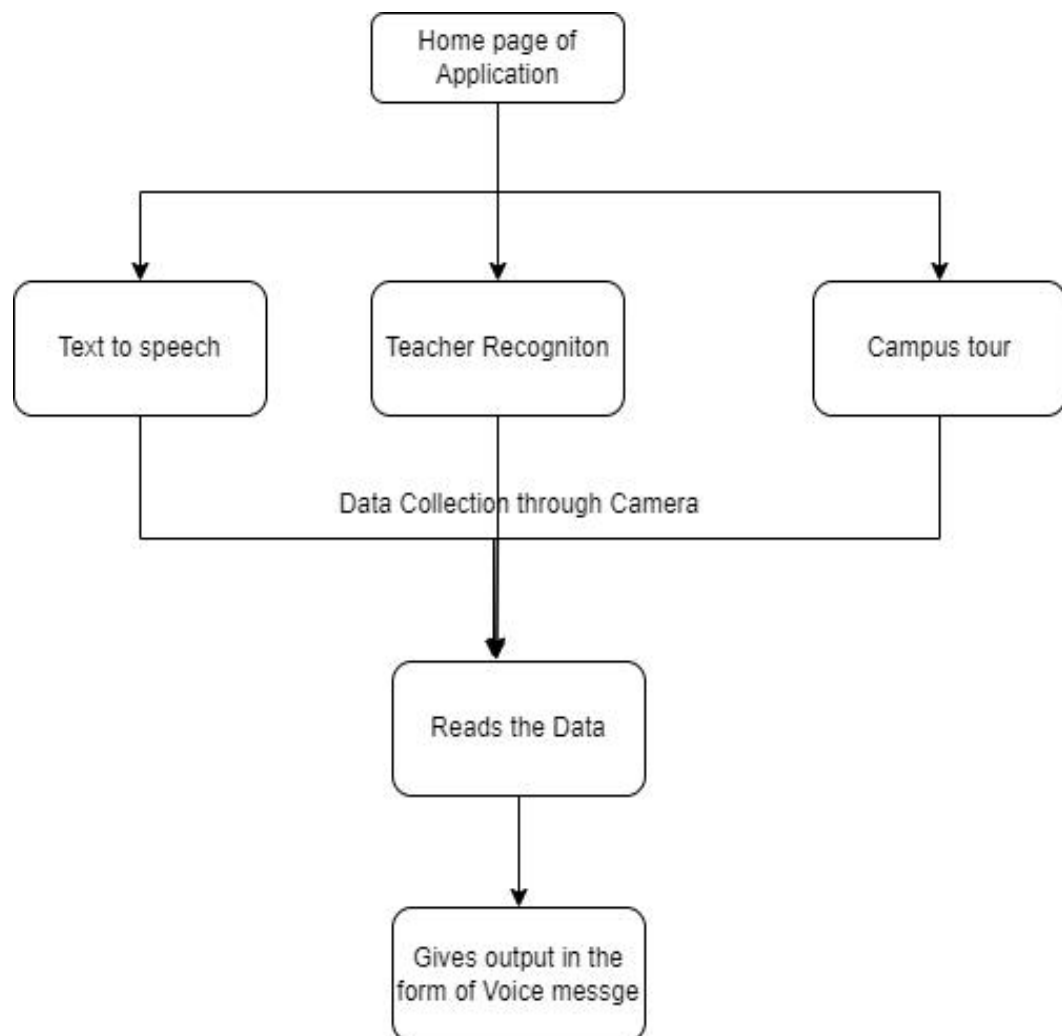- **Hard Disk:** 1 TB

Software Requirements:

- Operating System - Certified Distribution of WINDOWS
- Visual Studio
- Web Browser - Google Chrome
- Database(Backend) - SQL
- Python 3.6
- Tesseract
- Pytorch
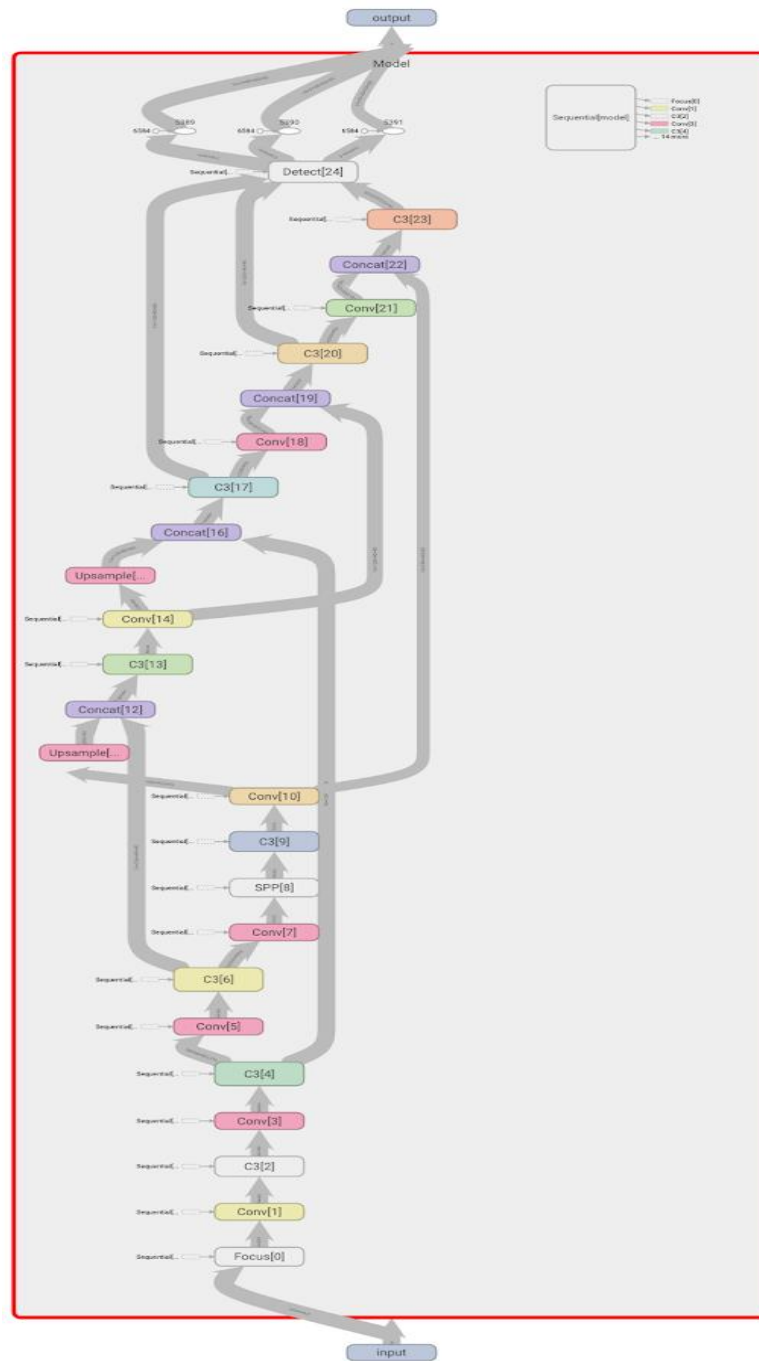- Scispacy

3. <u>DETAILED DESIGN</u>
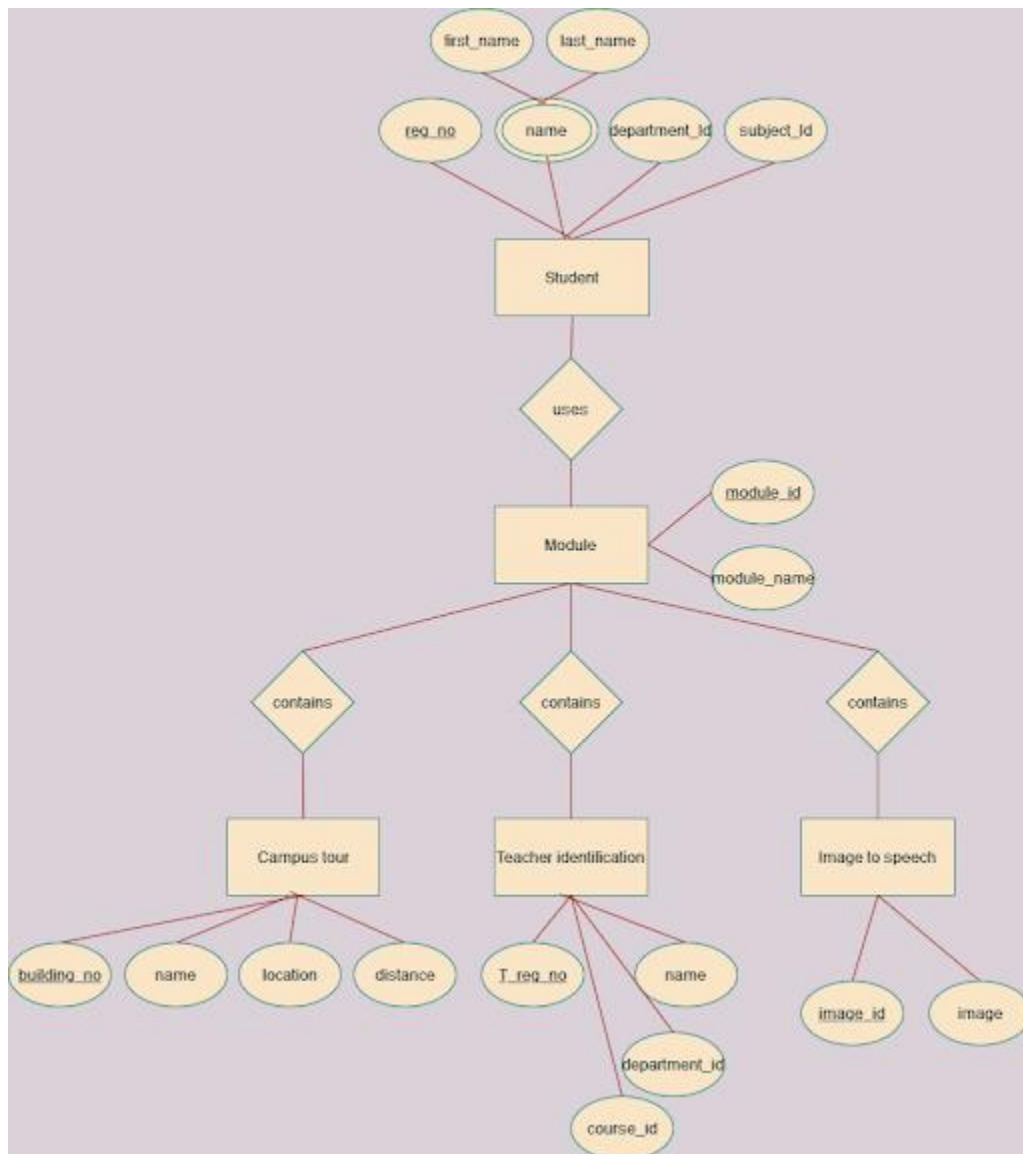
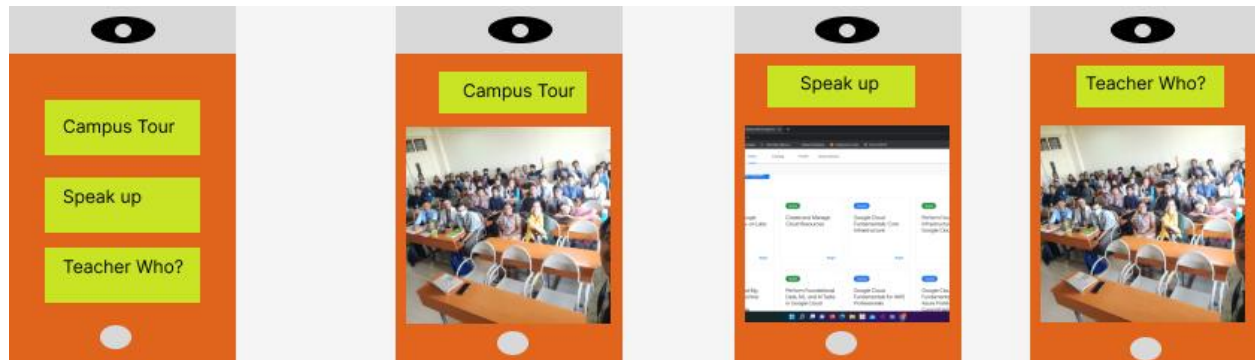## 3.1 <u>SOLUTION ARCHITECTURE</u>

3.2. FLOW DIAGRAM

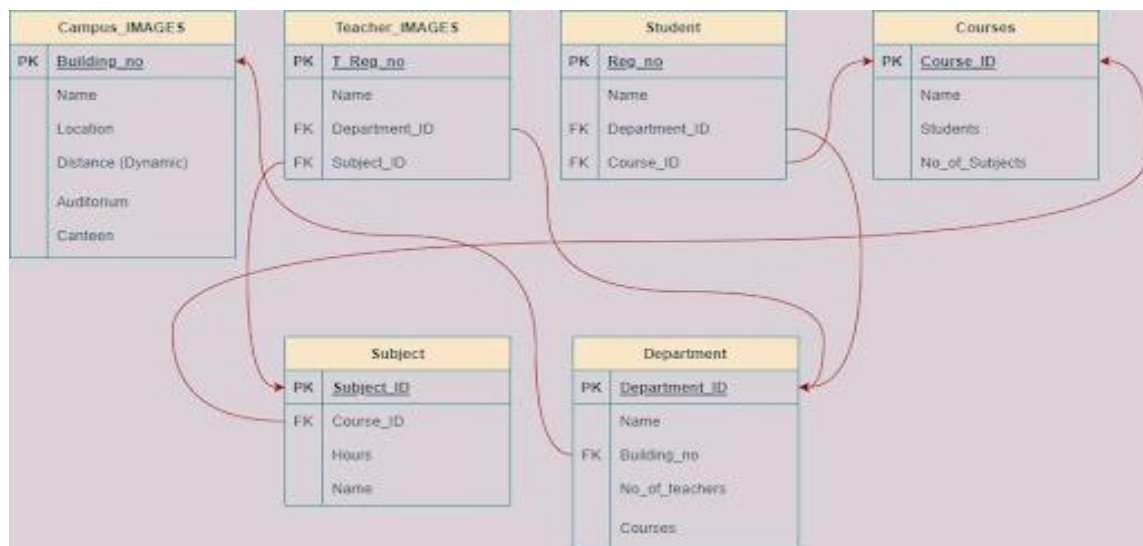## 3.3. ALGORITHM AND DATA STRUCTURE

3.4. ER DIAGRAM

3.5. INTERFACE DESIGN

3.6. CLASS DIAGRAM

4. PROJECT SPECIFICATIONS

4.1. SPECIFICATION CONCEPTS

1. **Image to Audio:** Text extraction is done using a Tesseract OCR tool from the uploaded image. The text from document images is detected and then using the Google text to speech recognition and python playsound the Image is converted into audio.

2. **Image Processing:** The images are trained using yolov5 and COCO128 and then the test images are feeded to check the models accuracy. We will also be using the functionalities in Android to access the camera to recognise the image. This includes the facial recognition and campus building recognition in our project.

4.2 MODULE DESCRIPTION

4.2.1. IMAGE TO SPEECH

When the user selects this module, then the camera will automatically be turned on and the camera will start detecting the images using the functionality of Tesseract OCR and the images with the text will automatically get recognised with the help of bounding boxes or annotated image box . The text in the image will be converted into text with the help of google text to speech and finally the recognised image's text will be voiced out with the help of python play sound.

4.2.2. TEACHER RECOGNITION

When the user selects this module, then the camera will automatically be turned on and the camera will start detecting the images. We have trained the model on YOLOv5 and COCO128 using these two deep learning algorithms. Now once the camera is pointed at any teacher this functionality will voice out the information about them.

4.2.3. <u>CAMPUS TOUR</u>

When the user selects this module, then the camera will automatically be turned on and the camera will start detecting the images. We have trained the model on YOLOv5 and COCO128 using these two deep learning algorithms. Now once the camera is pointed at any building in college then this functionality will voice out the information about them.

REFERENCES

-->Existing apps
    1. Seeing AI https://www.microsoft.com/en-us/ai/seeing-ai
    2. Be my eyes https://www.microsoft.com/en-us/ai/seeing-ai
    3. Look out https://blog.google/outreach-initiatives/accessibility/lookout-app-help-blind-and-visually-impaired-people-learn-about-their-surroundings/
    4. Aira: https://aira.io/
-->Articles and documents
    5. https://docs.ultralytics.com/
    6. https://tesseract-ocr.github.io/
    7. https://gtts.readthedocs.io/
    8. https://www.kaggle.com/ultralytics/coco128
    9. https://www.syndicai.co/blog/train-yolov5
-->Videos
    10. https://www.youtube.com/watch?v=oXlwWbU8l2o