

Realtime Object Detection and Disease Prediction of Visually Impaired People

Dr. Manohar Chaudhary¹, Pratik Gade², Swapnil Waghmare³, Vaibhav Misal⁴, Abhishek Waghmare⁵

Professor, Department of Computer Engineering¹

Students, Department of Computer Engineering^{2,3,4,5}

Sinhgad Institute of Technology and Science, Kusgaon(BK), Pune, India

Abstract: In order to detect objects, this research combines real-time object identification with appropriate deep learning techniques. In our study, we describe the creation of a real-time system for item recognition, classification, and position estimation in the open environment. Blind and visually impaired people daily deal with a variety of difficulties. The proposed plan's goal. Computer vision is used to precisely identify indoor items. Those who are blind or visually challenged can use navigational aids, The Technology for navigation of the blind is not sufficiently accessible, without vision it can be challenging for visually impaired persons to navigate through rooms or different road paths. The main aim to develop the project is to help the visually impaired people and to detect the obstacles. The blind persons life become easier and without anyone helps they can walk alone through street they does not need anyone to assist them they can handle their self correctly. The preventing users from dangerous location our aim is to collected from environment (cameras, sensors, scanners, etc.) and transmitted to the users to the audio format. Data in the healthcare industry consists of all the information related to patients. Herea general architecture has been proposed for predicting the disease in the healthcare industry.

Keywords: YOLO V3, Firebase, CNN, Realtime Object Detection

I. INTRODUCTION

1.1 Overview

Visually Impaired Persons faces many problem in general life in navigation and locating the objects.

For Better understanding we have to built some type of Realtime object detection technology where we can detect the object. Also the disease prediction functionality also gives the more functionality to this project.

1.2 Motivation

Motivation Enhance the lives of the visually impaired people through the creation of a wearable tool locate near and distant obstacles and person can easily understand about that obstacles in the form of audio.

Nowadays, humans face various diseases due to the current environmental condition and their living habits. The identification and prediction of such diseases at their earlier stages are much important, so as to prevent the extremity of it.

1.3 Problem Statement

The main goal of this project is to guide the blind people through the output of processor or controller by voice to navigate them.

This project exploring the possibility of using the hearing sense to understand visual objects. The sense of sight and hearing sense share a striking similarity both visual object and audio sound can be spatially localized

1.4 Objective

While the entire aim of this project is identify the objects and identifying the diseases using symptoms.

Following are the objectives:

- Measuring distance of object from camera .

- Measuring dimensions of an object from known distance .
- Identifying the diseases using Symptoms.
- Using artificial intelligence and the YOLOv3 Algorithm, detect an object and measure its distance and dimensions.

1.5 Project Scope and Limitations

The same object can have completely different shapes and sizes. Computer vision needs to do a lot of research to read an object and understand what it means. When it comes to video, detectors need to be trained to perform analysis in an ever-changing environment. One of the biggest difficulties of object detection is that an object viewed from different angles may look completely different.

1.6 Methodologies of Problem Solving

LR algorithm: Linear Regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

In practice, the linear regression algorithm analyzes relationships between variables. It assigns probabilities to discrete outcomes using the Sigmoid function, which converts numerical results into an expression of probability between 0 and 1.0.

II. LITERATURE SURVEY

2.1 Study of Research Paper

Paper No.1

Title. Object Recognition App for Visually Impaired

Author: Sumitra A. Jakhete, Avanti Dorle. Piyush Pimplikar

Year Of Publication: 2019

Keywords—Blind, Object, Detection, SSD, Tensorflow, TextToSpeech

Method of Algorithms Used: CNN(Convolutional Neural Network)

It integrates various techniques to build a rich android application that will not only recognize objects around visually impaired people in real time but also give an audio output to assist them as quickly as possible.

Bounding boxes are generated which predicts the certainty called as confidence score. This score lets us know that the bounding box consists of some object. The model is integrated into Android Application. The app uses rear camera of the smartphone for real time processing.

We gathered the COCO 2014 data for various objects including 80 classes. We also tried SSD implementation of which we achieved better accuracy and performance than YOLO. Next we were successful in developing a basic prototype in python language and which could detect and recognize objects using webcam, and provide voice output as well.

Conclusion:

Visually impaired people today can read using Braille script but it is still tough for them to recognise and interact with household objects and also on roads. In this project we developed an android application to aid the visually impaired people. It sends them an audio of the label based on the confidence score of the predicted object.

Paper No.2

Title :- Disease Risk Prediction by Using Convolutional Neural Network.

Author : Sayali Ambekar, Rashmi Phalnikar

Year Of Publication: 2019

Keyword: Data Mining, Heart Disease Prediction, Naïve Bayes, KNN, Heart disease risk prediction, CNN-UDRP algorithm.

Method of Algorithms Used: KNN Algorithm

The main idea of this paper is to predict whether the patient suffers from heart disease or not. And also predicting the risk of heart disease that is patient it is at high risk or low risk. The user enters the appropriate input values from his/her

health report. After this, the historical dataset is purposefully uploaded. The data imputation and data cleaning step is necessary. Then the naïve bayes and KNN algorithm is implemented on the input values and the heart disease is predicted

Conclusion

In this paper, we experiment the CNN-UDRP algorithm using structured data for disease risk prediction. We performed heart disease prediction using naïve bayes algorithm and KNN algorithm. In future, we will add more diseases and predict the risk which patient suffers from specific disease..

Paper No.3

Title:- Automatic Method for Measuring Object Size Using 3D Camera

Author:- Cuong Vo-Le, Pham Van Muoi, Nguyen Hong Son, Nguyen Van San, Vu Khac Duong and Nguyen Thi Huyen

Year Of Publication: 2021

Keyword: Size measurement, 3D camera, automatic measurement, key point extraction, depth interpolation

Method of Algorithms Used: Gesture recognition. Kinect.Yoga human computer interaction. supervised learning . CNN(Convolutional Neural Network).

The proposed system consists of two parts which are object detection and object measurement. The camera will capture a frame and the frame will convert to grayscale to increase quickness and accuracy. Object detection and measurement are the other two parts of the proposed system.

Conclusion

In our study, an accurate automatic method for measuring object size is proposed. The object detection is combined with our two proposed algorithms to extract key points and interpolate depth information. with the 3D coordinate of key points, the width and height of each object are calculated. Various experiments have been done to prove our proposed system can achieve accurate object size. our method forms the basic ideal to develop systems that can be applied to automation, logistics, and measurement in daily life.

III. SOFTWARE REQUIREMENTS SPECIFICATIONS

3.1 Introduction

3.1.1 Project Scope

In order to detect objects, this research combines real-time object identification with appropriate deep learning techniques. In our study, we describe the creation of a real-time system for item recognition, classification, and position estimation in the open environment. Blind and visually impaired people daily deal with a variety of difficulties. The proposed plan's goal. Computer vision is used to precisely identify indoor items. Those who are blind or visually challenged can use navigational aids.

3.1.2 User Classes and characteristics

After logging into our system, the user can access the GUI page. When camera is open the object is detected and audio response is created.

3.1.3 Assumptions and Dependencies

Assumption: In this system, we take an image dataset as input and, using the YOLOV3 algorithm, determine whether a object and measure distance and dimension. YOLOV3 algorithm:

YOLOv3 (You Only Look Once, Version 3) is a real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. YOLO uses features learned by a deep convolutional neural network to detect an object.

Dependencies:

Used Python Language: Python is commonly used for developing websites and soft- ware, task automation, data analysis, and data visualization. Since it's relatively easy to learn, Python has been adopted by many non-programmers such as accountants and scientists, for a variety of everyday tasks, like organizing finances.

Copyright to IJAR SCT

www.ijarsct.co.in

DOI: 10.48175/IJAR SCT-9547



90

Python is a general-purpose programming language, so it can be used for many things. Python is used for web development, AI, machine learning, operating systems, mobile application development, and video games. Python is a relatively easy programming language to learn and follows an organized structure

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention

3.2 Functional Requirements

3.2.1 System Feature 1 (Functional Requirement)

Proposed system consists of 4 modules:

- a) Feature point extraction: Feature points of each Dataset parameters gets detected.
- b) Feature correspondence matching: Matching of selected feature points across various parameters.
- c) Point estimation: Position estimation and vision system orientation during navigation.
- d) Position refinement: Location estimate based, accurate location derivation

3.2.2 System Feature 2 (Functional Requirement)

In system we have used YOLO v3 algorithm of Machine Learning.

3.3 External Interface Requirements

3.3.1 User Interfaces

User can login our system and then run App

3.3.2 Hardware Interfaces and Software Interfaces

RAM:8GB

As we are using Machine Learning Algorithm and Various High Level Libraries Laptop

RAM minimum required is 8 GB.

Hard Disk : 256GB

Data Set of CT Scan images is to be used hence minimum 40 GB Hard Disk memory is required.

Processor : M1 Apple Chip

IDE : Spyder

Spyder is a free and open source scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package. Coding Language : Python Version 3.8

Highly specified Programming Language for Machine Learning because of availability of High Performance Libraries.

Android Studio: Android Studio is an opensource application development environment for creating application

Google Play Console: it is a Application Development Publishing Environment Created by google for Play Store.

3.4 Non Functional Requirements

3.4.1 Performance Requirements

The performance of the functions and every module must be well. The overall performance of the software will enable the users to work decently. Performance of encryption of data should be fast. Performance of the providing virtual environment should be fast Safety Requirement.

The application is designed in modules where errors can be detected and steadily. This makes it easier to install and update new functionality if required.

3.4.2 Safety Requirement

The application is designed in modules where errors can be detected and fixed easily. This makes it easier to install and update new functionality if required.

3.4.3 Software Quality Attributes

Our software has many quality attribute that are given below:

- Adaptability: This software is adaptable by all users.
- Availability: This software is freely available to all users. The availability of the software is easy for everyone.
- Maintainability: After the deployment of the project if any error occurs then it can be easily maintained by the software developer.
- Reliability: The performance of the software is better which will increase the reliability of the Software.
- User Friendliness: Since, the software is a GUI application; the output generated is much user friendly in its behavior.
- Integrity: Integrity refers to the extent to which access to software or data by unauthorized persons can be controlled.
- Security: Users are authenticated using many security phases so reliable security is provided.
- Testability: The software will be tested considering all the aspects.

3.5 System Requirements

3.5.1 Database Requirements

DBSqlite 3: The Database Requirements involves the use of a lot of information, some which will be needed several times and the most appropriate form of storage of this data is in a database.

This will allow data to be saved from input to the Database Requirements and retrieved to be used by the Database Requirements. As an important aspect of this project is use of Time Control System. In this section several databases are reviewed for their suitability to this project.

3.5.2 Software Requirements

IDE: Android Studio

Coding Language: Java

Operating System: Mac Os

3.5.3 Hardware Requirements

Speed: 1.1 GHz

Hard Disk: 256 GB

RAM: 8 GB

Processor: Apple M1 Chip

3.6 Analysis Models: SDLC Model to be Applied

SDLC Models stands for Software Development Life Cycle Models. In this article, we explore the most widely used SDLC methodologies such as Agile ... Each software development life cycle model starts with the analysis, in which the Also, here are defined the technologies used in the project, team load.

One of the basic notions of the software development process is SDLC models which stands for Software Development Life Cycle models. SDLC - is a continuous process, which starts from the moment, when it's made a decision to launch the project, and it ends at the moment of its full remove from the exploitation. There is no one single SDLC model. They are divided into main groups, each with its features and weaknesses.

Waterfall model

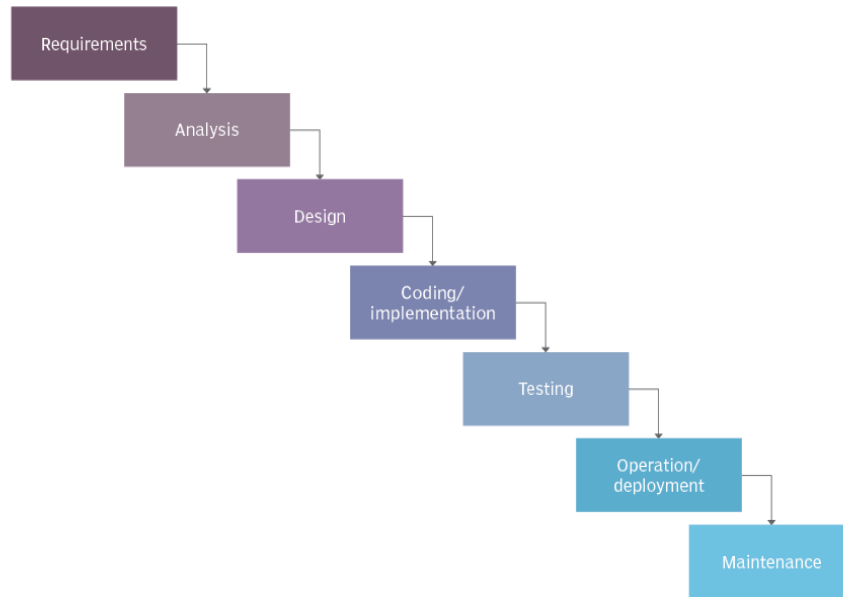


Figure: Waterfall Model

3.7 System Implementation Plan

The System Implementation plan table, shows the overall schedule of tasks compilation and time duration required for each task.

| Sr. No. | Name/Title | Start Date | End Date |
|---------|--|------------|----------|
| 1 | Preliminary Survey | 2-9-22 | 16-9-22 |
| 2 | Introduction and Problem Statement | 17-9-22 | 22-9-22 |
| 3 | Literature Survey | 24-9-22 | 3-10-22 |
| 4 | Project Statement | 4-10-22 | 6-10-22 |
| 5 | Software Requirement And Specification | 7-10-22 | 10-10-22 |
| 6 | System Design | 11-10-22 | 6-11-22 |
| 7 | Partial Report Submission | 20-11-22 | 28-11-22 |
| 8 | Architecture Design | | |
| 9 | Implementation | | |
| 10 | Deployment | | |
| 11 | Testing | | |
| 12 | Paper Publish | | |
| 13 | Report Submission | | |

IV. SYSTEM DESIGN

System Architecture

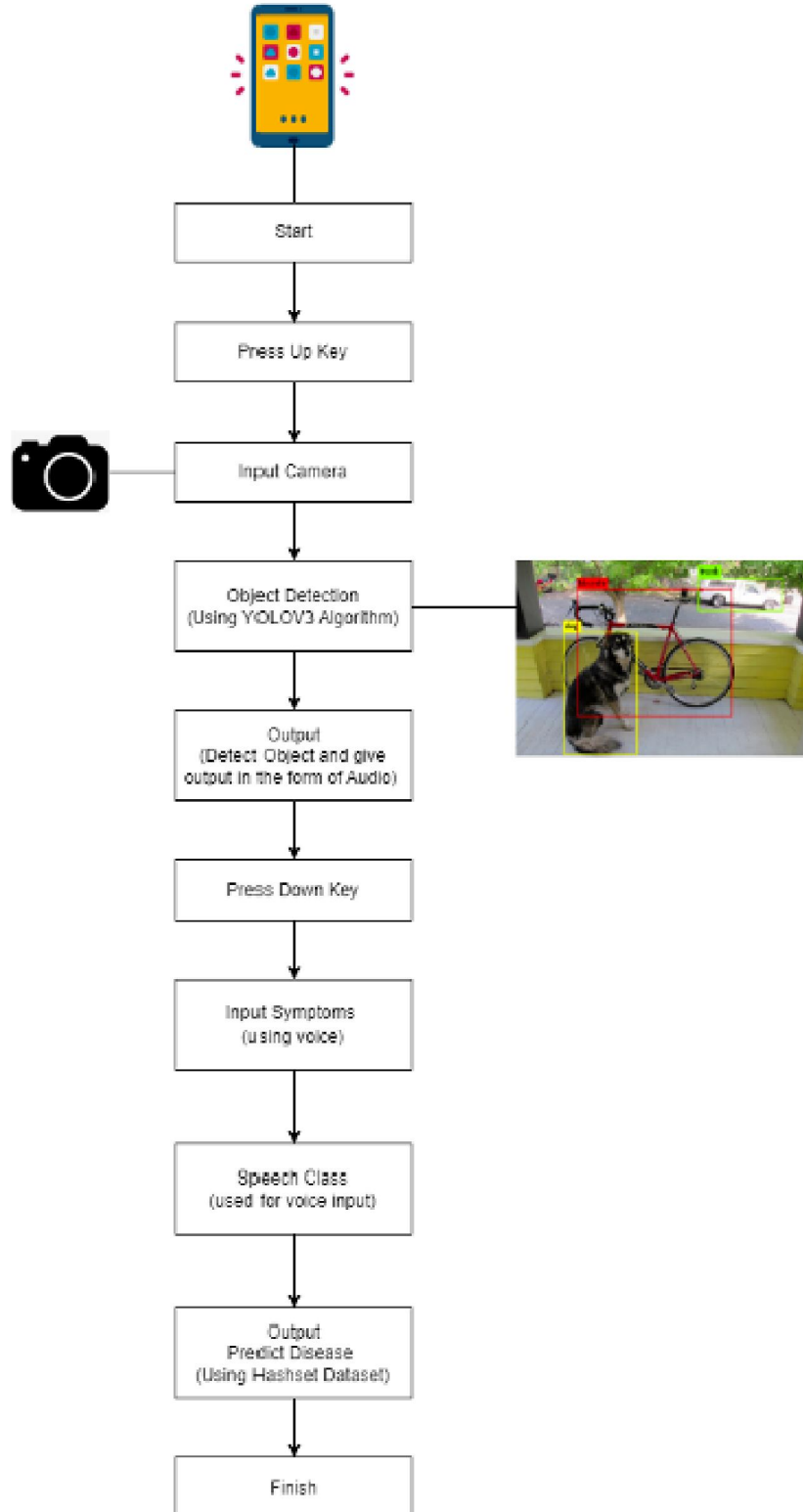


Figure: System Architecture

Data Flow Diagrams

In Data Flow Diagram, We Show that flow of data in our system in DFD0 we show that base DFD in which rectangle present input as well as output and circle show our system, In DFD 1 we show actual input and actual output of system input of our system is text or image and output is rumor detected likewise in DFD 2 we present operation of user as well as admin.

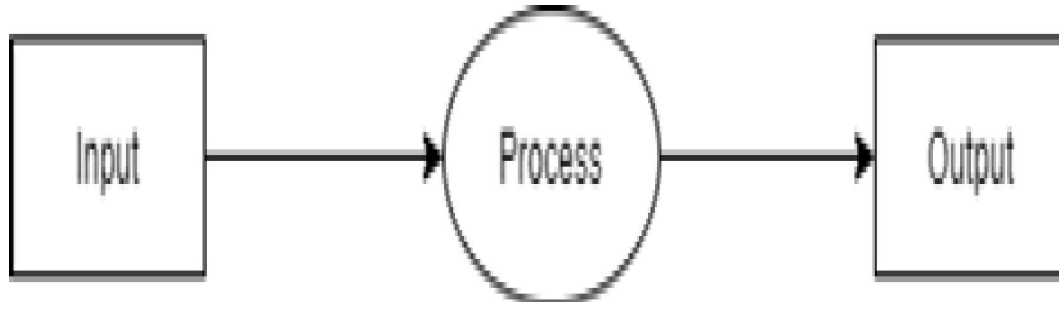


Figure: Data Flow Diagram 1

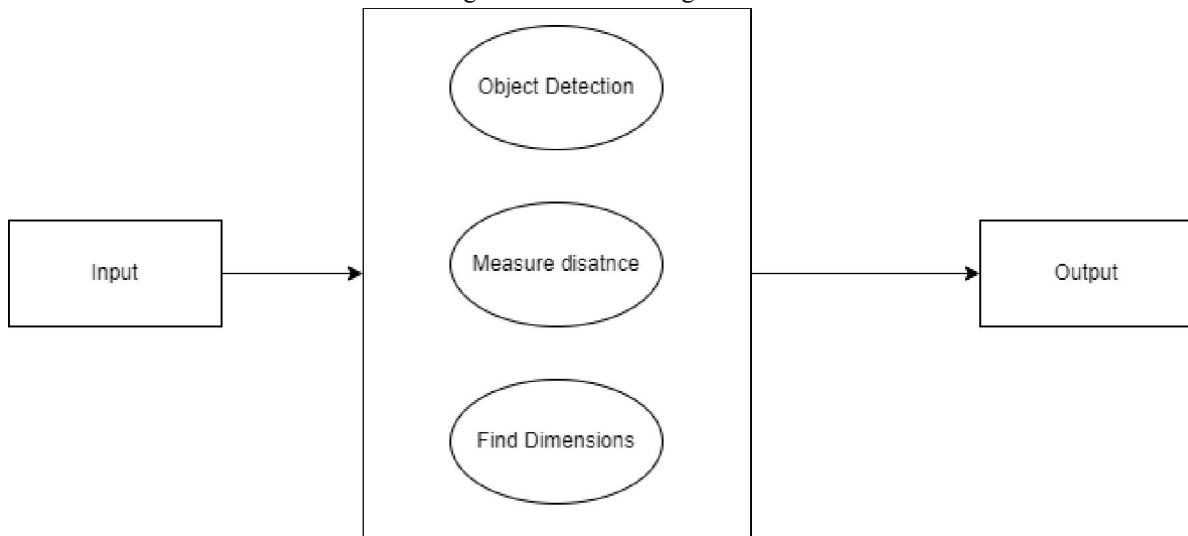
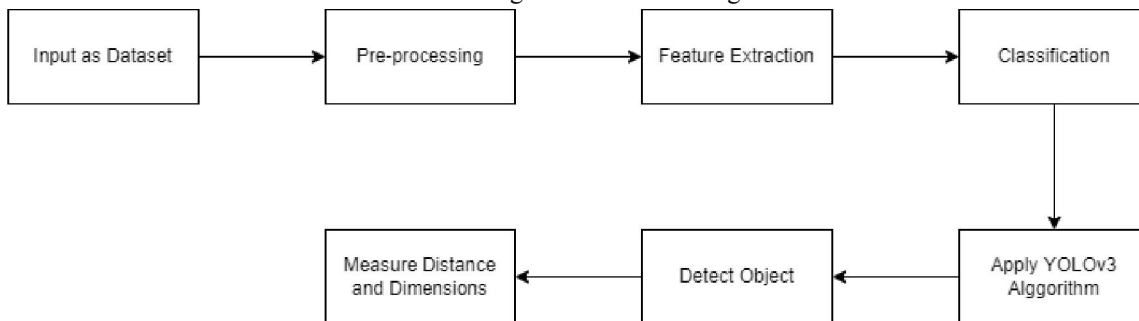


Figure: Data Flow Diagram 2



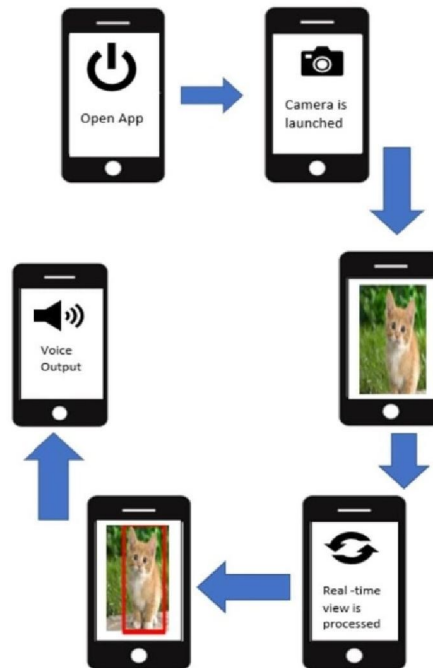
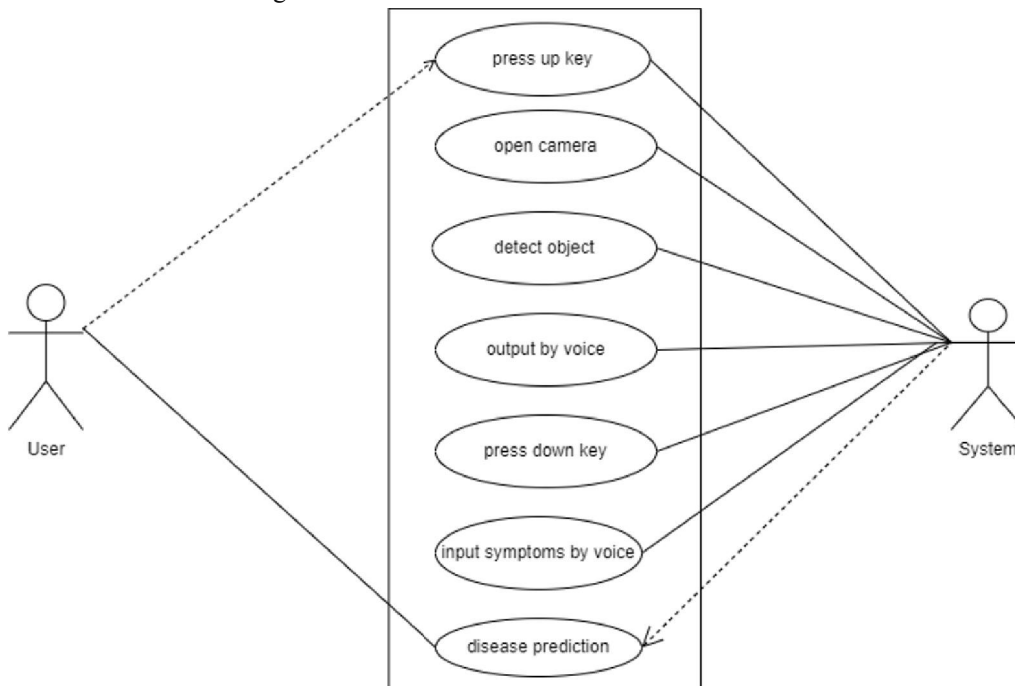


Figure: Data Flow Diagram 3

UML Diagrams

Unified Modeling Language is a standard language for writing software blueprints. The UML may be used to visualize, specify, construct and document the artifacts of a software intensive system. UML is process independent, although optimally it should be used in process that is use case driven, architecture-centric, iterative, and incremental.

The number of UML Diagrams are available:



Use Case Diagram

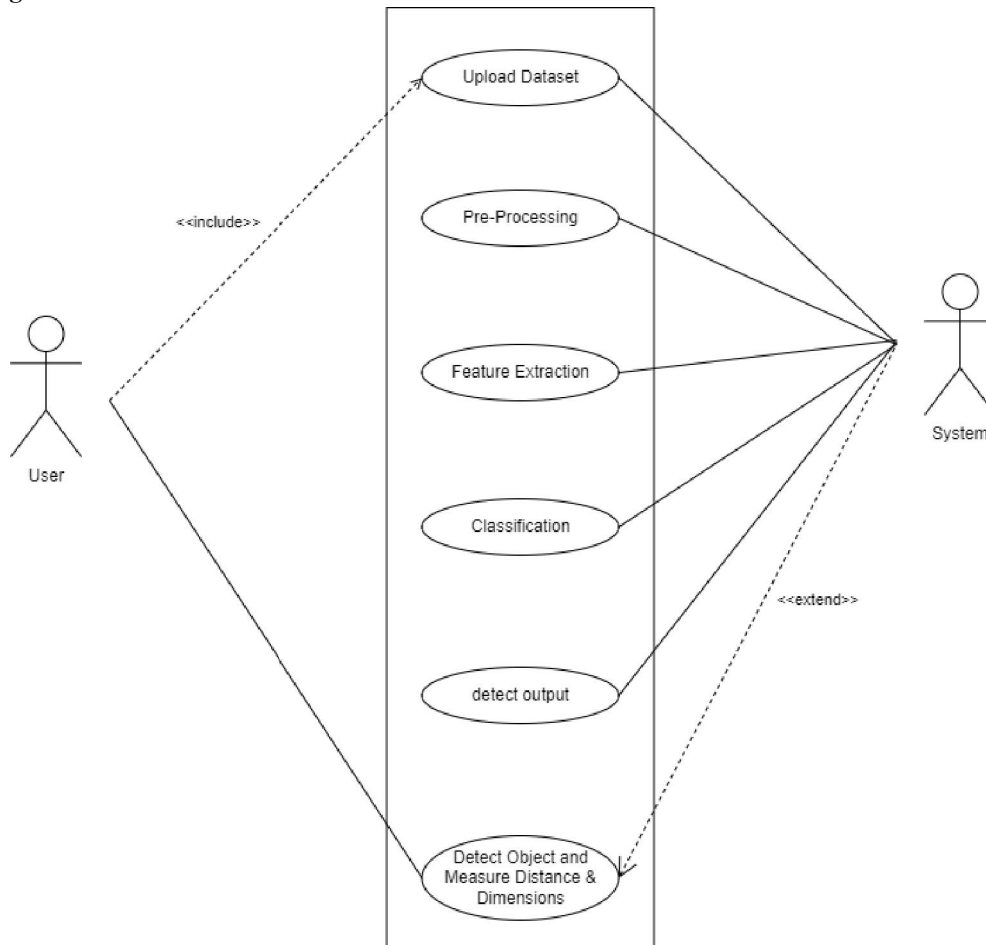


Figure: Use Case Diagram

2 Activity Diagram

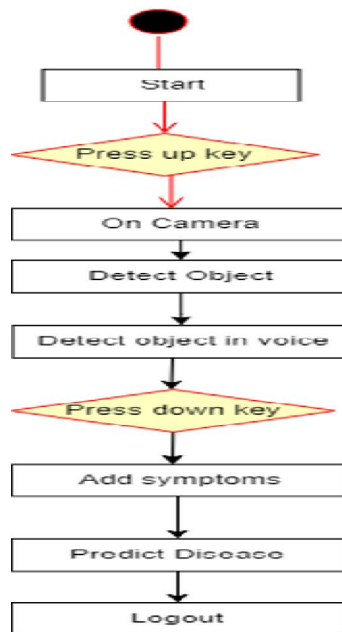


Figure: Activity Diagram

Class Diagram 1:

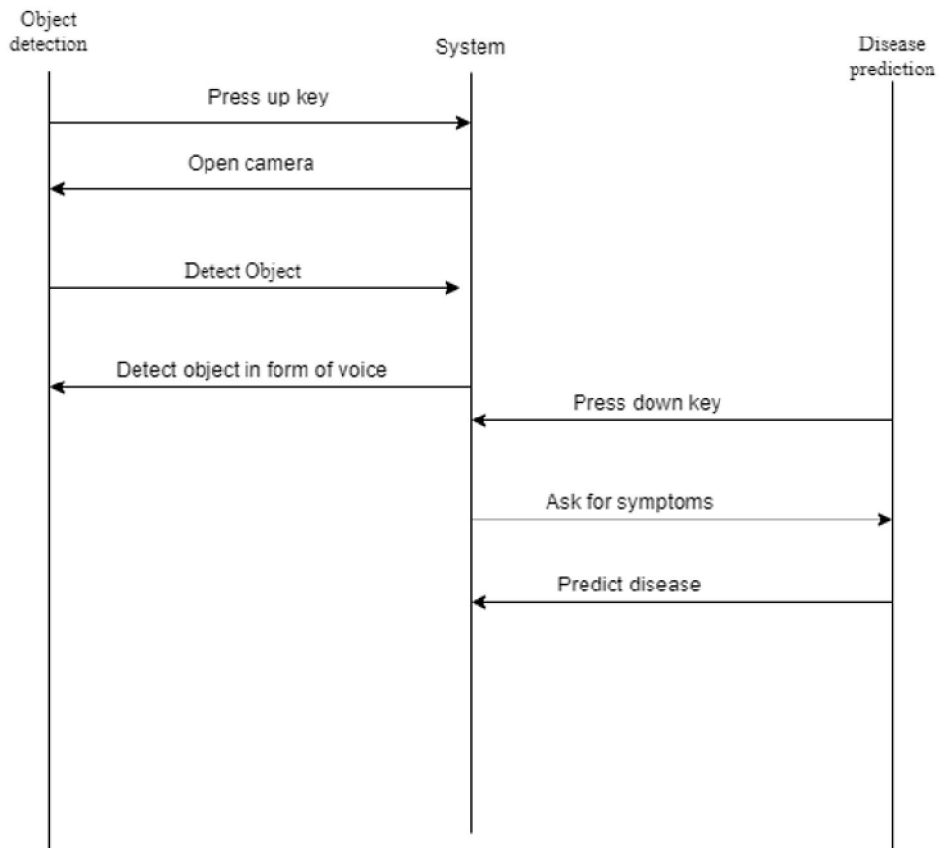


Figure: Class Diagram

Class Diagram 2

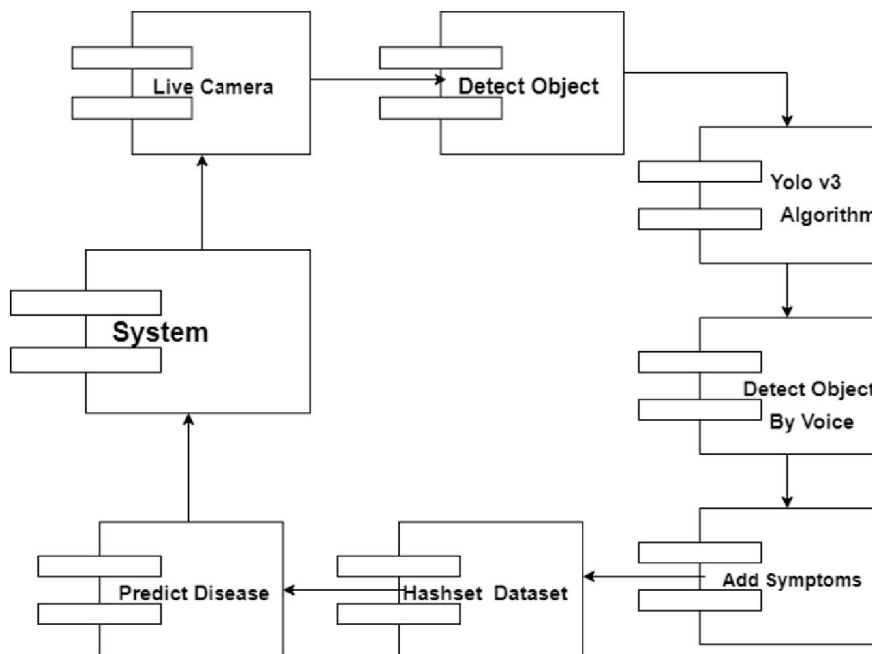


Figure: Class Diagram

V. ALGORITHM USED

5.1 TFTlite Module

The TensorFlow Lite module for object detection is a software library developed by Google that enables the creation of computer vision models for mobile and embedded devices. Specifically, it provides a set of pre-trained machine learning models that can be used for real-time object detection tasks, such as identifying and labeling objects within a live camera feed.

This module is particularly useful for visually impaired people, as it can be integrated into mobile apps and used to provide real-time object recognition and labeling, which can assist with navigation, object identification, and other tasks.

The TensorFlow Lite module for object detection is built on top of Google's TensorFlow framework, which is a popular open-source machine learning library. It uses a technique called convolutional neural networks (CNNs) to analyze images and identify objects within them.

Overall, the TensorFlow Lite module for object detection is a powerful tool for developing mobile apps that can assist visually impaired people with a wide range of tasks, from navigating unfamiliar environments to identifying objects around them.

5.2 Module Use Cases

Assistive technology for the visually impaired: As mentioned earlier, the module can be integrated into mobile apps to provide real-time object recognition and labeling, which can assist visually impaired people with navigation and object identification.

1. Smart home automation: The module can be used to detect objects and people within a home, allowing for more sophisticated automation of lighting, temperature control, and security systems.
2. Retail and inventory management: The module can be used to track inventory levels, monitor product placement, and detect customer behavior within a retail environment.
3. Industrial automation: The module can be used to monitor and detect objects within a manufacturing environment, allowing for more efficient and effective quality control.
4. Medical imaging: The module can be used to analyze medical images, such as X-rays or CT scans, to detect and diagnose conditions more accurately and quickly.
5. Robotics: The module can be used to enable robots to detect and interact with objects within their environment, allowing for more sophisticated and flexible automation in manufacturing and other industries.

Overall, the TensorFlow Lite module for object detection has a wide range of applications across various industries and use cases, making it a powerful tool for developing computer vision applications.

5.3 TFTlite Module Internal Implementation

The TensorFlow Lite module for object detection uses a type of neural network called a convolutional neural network (CNN) for object detection. More specifically, it uses a type of CNN called a Single Shot Detector (SSD).

The SSD architecture is designed to be fast and accurate for real-time object detection tasks. It works by dividing an input image into a grid of smaller regions and predicting the presence and location of objects within each region. The network is trained on a large dataset of labeled images to learn how to recognize different objects and classify them into categories.

During inference, the input image is first passed through a pre-processing step to prepare it for analysis. Then, the network processes the image in parallel across multiple layers to detect objects at different scales and aspect ratios. Finally, the network outputs a set of bounding boxes and class labels for each detected object.

The TensorFlow Lite module for object detection comes with pre-trained models that can recognize hundreds of different object categories, such as people, vehicles, animals, and household objects. These models can be fine-tuned or retrained on custom datasets to recognize specific objects or categories for specific use cases.

Code: Note that this is just a simple example, and the actual implementation may vary depending on the specific use case and requirements.

```
import tensorflow as tf
```

```
import numpy as np
import cv2
# Load the TFLite model and allocate tensors
interpreter = tf.lite.Interpreter(model_path="model.tflite")
interpreter.allocate_tensors()
# Get input and output tensors
input_details = interpreter.get_input_details()
output_details = interpreter.get_output_details()

# Load image
image = cv2.imread("image.jpg")

# Preprocess image
image = cv2.resize(image, (300, 300))
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
image = np.expand_dims(image, axis=0)
image = (2.0 / 255.0) * image - 1.0

# Run inference
interpreter.set_tensor(input_details[0]['index'], image)
interpreter.invoke()
boxes = interpreter.get_tensor(output_details[0]['index'])
classes = interpreter.get_tensor(output_details[1]['index'])
scores = interpreter.get_tensor(output_details[2]['index'])
num_detections = interpreter.get_tensor(output_details[3]['index'])

# Print the results
for i in range(int(num_detections)):
    if scores[0][i] >= 0.5:
        ymin = int(boxes[0][i][0] * image.shape[1])
        xmin = int(boxes[0][i][1] * image.shape[2])
        ymax = int(boxes[0][i][2] * image.shape[1])
        xmax = int(boxes[0][i][3] * image.shape[2])
        cv2.rectangle(image, (xmin, ymin), (xmax, ymax), (0, 255, 0), 2)
        cv2.putText(image, str(classes[0][i]), (xmin, ymin - 10), cv2.FONT_HERSHEY_SIMPLEX, 0.5, (0, 255, 0), 2)

# Show the image
cv2.imshow("Object Detection", image)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

In this code, we first load the TFLite model and allocate tensors for input and output. We then load an image, preprocess it to the required size and format, and run inference using the interpreter. Finally, we extract the output tensors, filter the results based on a score threshold, and draw bounding boxes and class labels on the image.

VI. CONCLUSION AND FUTURE WORK

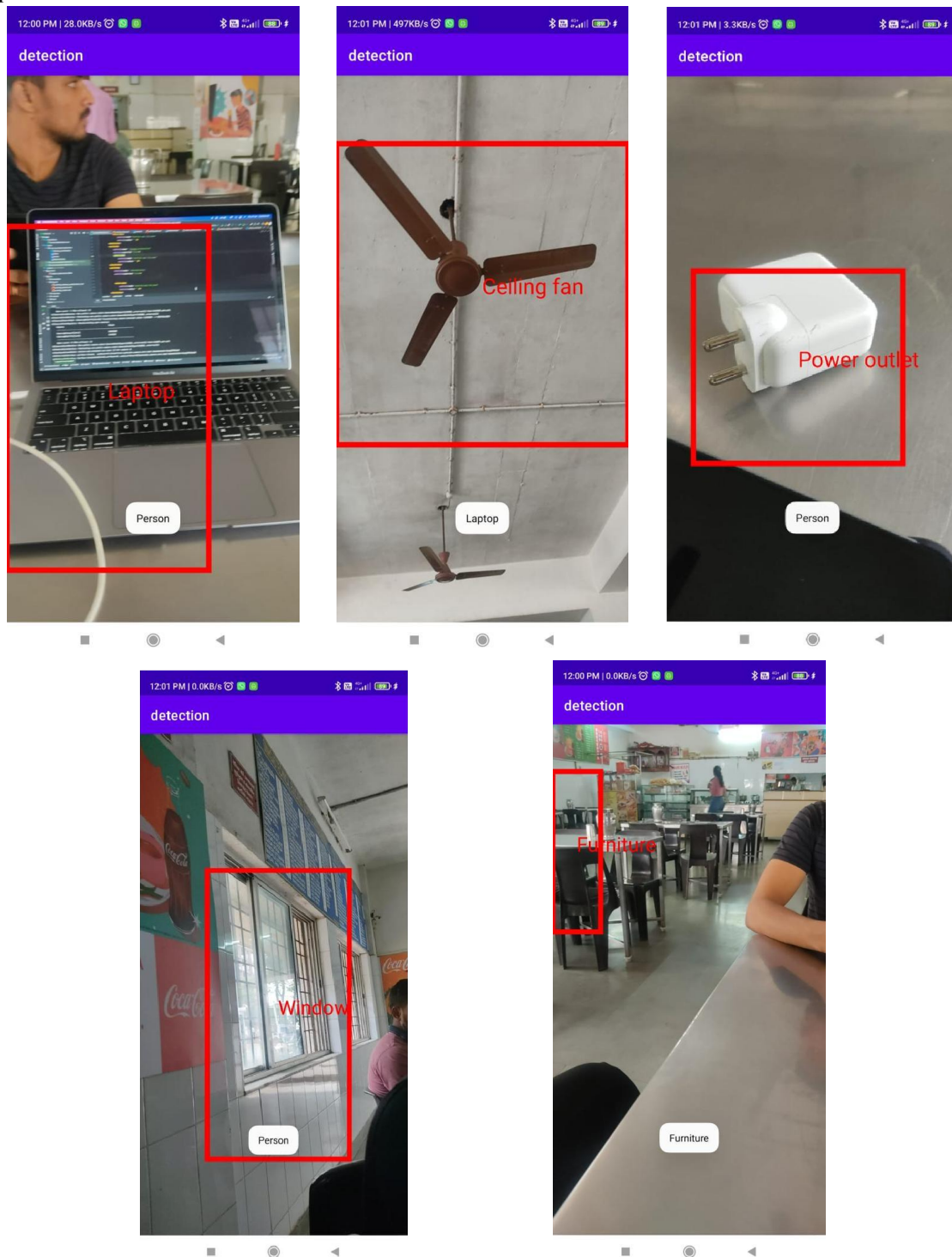
6.1. Conclusion

The application has a very simple and easily navigable User Interface that suits the visually impaired users. As soon as the application is launched, the camera will start capturing the real time video. As soon as the user presses a button, the

server-side backend algorithm will start processing it and notify the user accordingly as output audio. The Yolo algorithm can be stopped by pressing the same button again. This is how objects around the blind people and their positions are detected and conveyed to them via an audio output using the YOLOv3-tiny algorithm.

Proposed a system to predict the disease based on symptoms given by user. We provide Hash-set dataset for disease prediction.

6.2. Experimental results



VII. REFERENCES

- [1]. Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, "Recent advances in convolutional neural network acceleration," *Neurocomputing*, vol. 323, pp. 37–51, 2019
- [2]. Nasreen, W. Arif, A. A. Shaikh, Y. Muhammad and M. Abdullah, "Object Detection and Narrator for Visually Impaired People," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, 2019, pp. 1-4, doi: 10.1109/ICETAS48360.2019.9117405
- [3]. S. Vaidya, N. Shah, N. Shah and R. Shankarmani, "Real-Time Object Detection for Visually Challenged People," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 311- 316,doi:10.1109/ICICCS48265.2020.9121085
- [4]. V. Mohane and C. Gode, "Object recognition for blind people using portable camera," IEEE WCTFTR 2016 - Proc. 2016 World Conf. Future. Trends Res. Innov. Soc. Welf., pp. 3–6, 2016.
- [5]. H. Jabnoun, F. Benzarti, and H. Amiri, "Visual substitution system for blind people based on SIFT description," 6th Int. Conf. Soft Computer Pattern Recognition, SoCPaR 2014, pp. 300–305, 2015.