CS491 Senior Design Project I

Project Specifications Report

Spring 2020

Cemil Şişman	21502560
Derviş Mehmed Barutcu	21302589
A A M Jubaeid Hasan Chowdhury	21503356
Mustafa Azyoksul	21501426
Onur Mermer	21600878

1. Introduction

It is hard for blind people to navigate. Even indoors, they have a hard time recognizing their surroundings. With the advances of technological developments and image processing algorithms, these kinds of people can benefit from these advances.

Image processing is everywhere in our lives. Most of the phone manufacturers have been implementing image processing and machine learning techniques into their phone to provide better user experiences. However they are limited with the usage of the consumers. They are used to make photos more accurate or videos to look better or scan QR codes.

Our aim is to write a mobile application to help the visually impaired users. We want to use such technology that we use in normal lives to help the community.

1.1. Description

EyeSight is a mobile application that helps visually impaired people to be aware of their surroundings. EyeSight uses a rear camera and a speaker or an earphone, processes each frame that is recorded from the rear camera and analyzes the objects in the frame. After analyzing the objects in the frame, it gives constant stream of information to the user through speaker or earphone.

EyeSight is for indoor use only since there will be vital situations outdoors and this would not be appropriate considering the scope of this project. As indoor complexity, parameters and vitality are lower than the outdoor, it would be better to implement for indoor usage for our case due to the time constraints that we have.

1.2. Constraints

This section is about the constraints of the project in terms of economic, implementation, time, ethical and sustainability.

1.2.1. Economic Constraints

- The application will be free to use for everyone and will not require or offer any in-app purchases.
- The application might require users to have a phone with a strong proximity sensor attached to it. If some phone does not have a strong proximity sensor, the user might need to purchase and attach a strong proximity sensor to his/her phone.

1.2.2. Implementation Constraints

- The application will be a mobile application developed for Android. A web version isn't suitable for daily use for the people who have impaired vision therefore we will not implement a web application.
- We will be using image processing and object classification and text to speech synthesis.
- The major challenge is understanding the context after recognizing objects. For this purpose we will utilize differences between one frame to the next frame. If something moves or the user moves the frames will be different.
- The language will be mainly Java and some Python for image processing.
- We will use Google Cloud Vision for object detection [1] and Google Cloud Text-to-Speech [2] for speech synthesis.

1.2.3. Time Constraints

- We are planning on starting the implementation during the summer of 2020 and complete the project by December 2020.

1.2.4. Ethical Constraints

- The application revolves around a very serious matter. That's why we will try to be as accurate as possible. We will use as much data as possible from various sensors and use past results to give the best possible explanation of the current scenario.
- The app will not force the blind person to make certain decisions. It will only try to factually present information about his/her surroundings.

1.2.5. Sustainability Constraints

- There will be expected errors with classification and recognition of some objects. We will try to optimize the app using past data.
- The application will consume the battery at a high rate because the app requires the mobile phone's camera to be recording. However, if the electricity is coming from a renewable source then that's not a problem.

1.3. Professional and Ethical Issues

- Since the application could result in undesirable circumstances, it's extremely important to work precisely in order to minimize the margin of error. It would be a major blunder if a visually impaired user suffered from a situation that should have been prevented by the application.

2. Requirements

This section explains the functional and nonfunctional requirements of the project.

2.1. Functional Requirements

Functional requirements are listed below.

- The user's relatives can take a photo of their faces to introduce themselves to the application. After they take their faces and enter their names, each time they enter the camera frame, EyeSight will tell their names to the user.
- EyeSight will process the video inputs from the camera and identify the objects in the frame.
- EyeSight will tell the objects' name, their color and their position in the frame to the user.
- If the app detects an emergency situation or the blind person presses the emergency button, the app will automatically call the guardian's phone number.

2.2. Nonfunctional Requirements

Nonfunctional requirements are listed below.

2.2.1. Usability

- Since the target of the application is visually impaired people, the application should have features that help users vocally.
- Visually impaired users might need help from others to use some of the features of the application.

2.2.2. Scalability

- The app should be scalable to any number of users. The processing power of the backend image detection servers need to be scalable to any number of users.

2.2.3. Extensibility

- Since the outdoors might increase the user's chance of facing a vital circumstance, the scope of the application is initially aimed at indoors only. But it is always possible to extend the application to be usable outdoors.

2.2.4. Performance

- Taking the frame from the camera of the user, processing it and creating the voicemail shouldn't take more than 1 seconds. It might take a shorter or longer time depending on the complication of the scene.

2.2.5. Reliability

- Reliability is a major requirement for the application. It should always be working precisely most of the time to avoid any undesirable circumstances. The margin of error is very small since the errors might result in vital situations.

3. References

[1] https://cloud.google.com/vision

[2] https://cloud.google.com/text-to-speech