Tensorflow based mobile application for indoor object identification for visually impaired

H. L. P. Chathurika
2018

Tensorflow based mobile application for indoor object identification for visually impaired

A dissertation submitted for the Degree of Master of Computer Science

H. L. P. Chathurika
University of Colombo School of Computing
2018

Declaration

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.
To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Student Name: H. L. P. Chathurika
Registration Number: 2015/mcs/015
Index Number: 15440152

_____

Signature: Date:

This is to certify that this thesis is based on the work of
Mr./Ms.
under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:
Supervisor Name: Dr. G. D. S. P. Wimalaratne

_____

Signature: Date:

Abstract

Vision is considered as one of the primary senses of human in obtaining inputs from the neighboring environment. It plays a vital role in human perception in identifying the environment around him. People who are visually impaired face many challenges and difficulties while interacting with the surroundings in visualizing images and objects. Detecting and recognizing generic objects needs to be performed as accurately and efficiently as possible. Many tools and techniques have been identified as an assistant to those people and as a support in decreasing the complications in their lives. So, industry has created a variety of computer vision products and services by developing new electronic technologies for the visually impaired in order to overcome the difficulties.

Mobile computer vision is identified as a promising technology to support the people who are visually impaired. Fresh and exciting possibilities are being introduced with the emerging technology. Deep learning methods have resulted in significant performance improvements in several application domains and as such several software frameworks have been developed to facilitate their implementation. Object identification is considered as a major research area that has been focused in deep learning domain.

The proposed system is a research-basedapplication, which uses the advantage of Tensorflow deep learning framework. The frameworkcan be used in many scenarios andfor this system it's used to detect and identify indoor objects for visually impaired people. Thousand sets of indoor objects data are being fed to the system and used for the object identification.Audio guidance is also provided by the developed system to convey the response. The main idea of the project is to provide a better and accurate system in identifying the objects for the people who are visually impaired.The system is evaluated in two approaches. User acceptance testing is done in order to identify the obstacles and difficulties that the visually impaired people get when using the proposed system. Object detection evaluation is carried out in order to identify the efficiency and the accuracy of the system.

Acknowledgement

This dissertation arose as an effort of number of people whose contribution inassorted ways in the realization of this research deserves special mention. It is a pleasureto convey my gratitude to them all through this humble acknowledgment.

I gratefully acknowledge Dr. G. D. S. P. Wimalaratne, supervisor of the project, for hissupervision, advice, and guidance from the very early stage of this project. Above alland the most needed, he provided me unflinching encouragement and support invarious ways. His truly scientist intuition has made him as a constant oasis of ideasand passions in teaching, which exceptionally inspire and enrich our growth asstudents, and as professionals want to be. I am indebted to him more than he knows.

I would also acknowledge my parents and all thefriends who provided their support in making this research a success.

Finally big thank to each and every person for the support and the assistance given tomake this a success.

Table of Contents

List of Figures

List of Tables

Chapter 01

Introduction

1.1 Introduction

Vision is considered to be one of the most essential human senses. Good vision is considered as a precious gift. It plays a vital role in human's day-to-day life. Vision helps us to connect with our surroundings and lead us to a better quality life by providing awareness of the environment. Good eyesight can interpret the colours, shapes and dimensions of the objects surrounded by us and provide us the brief idea about what is there and what will happen next. It does not take much time to identify the difference between a car and a bus for a person whose vision is normal.

There are people who are not privileged or gifted like many others. People suffering from low vision or blindness are facing many problems engaging in their day-to-day life. Millions of people exist in the world that is in incapable of understanding the environment due to visually impairment [5]. Since vision is considered to be one of the most important organ to sense the surroundings, its loss can be significantly reduce the individual orientation and mobility.

According to the International Classification of Diseases, there are four levels of visual function [1]. Normal vision is the first level, which denotes that the user is having a good vision. And there is moderate visual impairment where the user is having minor problems in his vision. It can lead to severe visual impairment where the user has severe effect in the normal vision. And Blindness is considered to be the fourth visual function where the user is unable to see anything.

Moderate visual impairment combined with severe visual impairment can be grouped under the low vision category. All the low vision people and the blind people together are considered as the visually impaired. Based on a study done by World Health Organization (WHO) in August 2014, 285 million people are estimated to be visually impaired. Among these statistics, 39 million are considered to be blind and

other 246 million were suffering from low vision [1]. These numbers are predicted to significantly increase by the year 2020.

Few years back, being visually impaired meant being destined to a life full of restriction [14]. Those people are being considered as a burden to the family as well as the society. The challenges faced by blind people in their everyday lives are not well understood. The developmental learning, carriers and communication of the visually impaired people are affected by the problems they have to face in their lives. A deeper understanding of the questions that those people would like to ask may help in identifying the approaches that need to be taken in providing solutions [3].

One of the highly needed approaches they want is a proper way to identify objects around them since they cannot live a life, which is full of hazards and problems. This problem is increased when the person is in unfamiliar and complex indoor environments. When we consider a visually impaired person, they are facing huge number of difficulties in their day today life. For an example, if they are having a conversation with few other people, they do not know whether a person is talking to them or someone else during that conversation [4]. This is only a basic problem they face in their normal lifestyle.

1.2 Motivation

People who are visually impaired often have a difficult time in self-navigating through the environments, which are outside the well known. Travelling inside a large interior space is also considered to be a great difficult task for those people. Because of that, they may need the help of a family member or a friend to guide them to navigate in an unknown environment. Even inside a place they usually walk, they need every detail about the obstacles such as chairs, tables and they expect them to remain in one location and assume any other will not change them. Everyday objects that are present in most environments become real obstacles for the visually impaired people [4].

Even though these people can discover alternative approaches to deal with their daily routines, it has become a hazard to them in other situations. It is significantly difficult for them to find a particular room in an unfamiliar environment and large indoors spaces such as hotels, shopping malls [5]. If they need to use a washroom, they will not be able to figure out which one is the correct one. If they need to find a fire exit in an unexpected situation, how can they figure out which door they should use?

This has become a concern and a huge issue for the visually impaired people and the need of a proper guidance is considered as an essential desire to direct them. An independent life is an expectation of every person. People with visual impairments are destined to be dependents of other people. The dependency needs to be reduced or if possible eliminated from the society in order to make them feel that they are not a burden to the society.

With the use of emerging technologies, the possibility of providing a solution to the problems faced by the visually impaired people have become achievable. In earlier days, touch devices like mobile phones, tablets were assumed to be inaccessible by the visually impaired people. Nowadays, they are getting familiar with the technology and it can improve their lives by providing solutions based on those emerging technologies [3].

1.3 Statement of the problem

People who are either suffering from low vision or fully blind, face so many problems in their day today lives. When we think about the normal daily routine of a person, these people face lots of challenges. One of the main challenges that this demographic has to face in their everyday lives is orientation and mobility in both outdoor and indoor scenarios [6]. Even though they managed understanding the surroundings and the environments where they deal with their daily routines, they are incapable of identifying new environments and it takes time for them to get aware about the new places. They need the guidance of another person to get used to that new environment.

A grocery shopping is something that is not a huge problem to the normal users. The necessary items can be selected and purchased without a hazard. For the people who are visually impaired, this is considered to be a considerable concern. When they go grocery shopping, if they want to buy some soft drinks, they cannot identify which one is Pepsi and which one is Sprite. Typically they will have to call someone and acquire their assistant. Even though the visually impaired may be getting the guidance of another person, no one has committed his or her life to be the vision of another person. The others may be busy with their own lives, so that the visually impaired cannot always complain not having anyone around to help or guide them. They need to figure out a way where they can be independently figure out their own issues. No one is happy about being a dependent of another person.

Another problem they are facing is that although they get to know about the objects around them, they do not have any clue about exact awareness about the in detailed picture. For an example, if there is a mug kept on a table, we get to see the colour of the mug. If there are some wordings written on it, we get to see that one too. Or else if there is a printed picture on the mug, we can see it clearly and admire it. For visually impaired, it is just another mug which provides no additional details. They do not get to know the exact image of the mug in front of them. By improving their vision towards objects by providing all the little details around them, we can improve their quality of life. Millions of people are suffering from these kind of problems in visual impairments and it is significantly important to provide them a way to have an idea about what is there unseen by them.

What they need is an alternative guidance where they can figure things out on their own without depending on other people. With the emerging technology, this problem can be addressed in a way where the visually impaired can take the advantage of the mobile technology to solve their day-to-day problems. Although there are guidance systems available for such purposes, they have some drawbacks that hamper their direct application in real-life situations.

1.4 Aim and Objectives

This project aims to develop an effective response to the challenges currently faced by people who are visually impaired by providing a mobile-based approach for real time object identification in an indoor environment by providing feature extraction and guidance about their surroundings to make better and quality experience.

- Provide a mobile based application which provides object detection of the surroundings in an indoor environment
- Provide a detail feature recognition of the detected object
- Provide an accurate object identification system
- Provide audio based guidance in recognizing the objects
- Provide an easily manageable mobile application that can be used by the visually impaired

1.5 Scope of Project

This is a mobile-based application, which targets visually impaired people who are either fully or partially blind. From this project, we try to transform the visual world, which cannot be seen by those people into an audio world by providing a way to identify the objects around them. Even they get to know about what is in front of them, they do not get the real image that we get to see. From this project, we target to provide a mechanism for the visually impaired, to get the detailed description of what is around them in an indoor environment by providing feature extraction of the real objects.

For example, inside a building, there are few indoor landmarks to identify which ones are the washrooms, fire exits, rest rooms and other important signs. For a person who is visually impaired, it is hard to obtain an idea about what is in front of them. From this application, object detection and identification inside an indoor environment will be targeted and inform the visually impaired person about the surroundings.

For providing the solution, TensorFlow machine-learning library is going to be used, which is acknowledged as an enhanced machine learning expertise that can be run on mobile devices as well. It is an open source deep learning library recently released by Google along with Python APIs developed. Using that emerging technology, we need to maintain pre-trained models for identifying the objects around us. And they have come up with a solution based on neural network where the visual recognition is done close to human brain.

As the initial step, the objects enclosed around the person needs to be identified. For detecting the objects, camera of the user's phone is going to be used. When the user points the camera to a certain location, the objects in that area is going to be identified by this suggested application. Also this focuses on providing the exact feature recognition of the detected object by maintaining pre-trained models using the TensorFlow machine-learning library. The spatial location of the detected object can be calculated considering the camera point and the visually impaired will be given audio information about the objects detected.

1.6 Structure of the Dissertation

Introduction of the project "TensorFlow based mobile application for indoor object identification for visually impaired" is described in this chapter. In chapter 02, the background information of the project and information about similar approaches is described. Chapter 03 describes the designing phase of the project with the methodologies used. Chapter 04 provides a detailed explanation about the proposed solution details and the implementation of the project. Project evaluation criteria and the findings are described in the 05th Chapter. In chapter 06 further works is described along with the conclusion of the project. References are stated at the end of the dissertation, which is followed by the appendix.

Chapter 02

Background of the project

2.1 Introduction

Vision impairment can be considered as a major issue that the humans face in their day-to-day lives. According to World Health Organization, the number of visually impaired people has been increased with the time. Visual impairment can be categorized as fully blind and half blind. A main challenge that these people face can be identified as the issues that they face in identifying objects in their environment. Accurate object identification is considered to be essential to provide a better quality life for the people who are visually impaired.

Even though many systems have been discovered, a satisfied approach is not yet introduced. It's highly essential to develop a system, which facilitates better and accurate object identification. For that, entire idea about the other prevailing systems has to be concerned and referred in detail. This chapter focuses on identifying the most important prevailing systems addressing the same problem and their features and limitations.

2.1 Object identification system

Uncorrected refractive errors and un-operated cataract are identified as the top two causes of vision impairment [1]. Visual impairment presents severe consequences in several capabilities related to visual functions in normal day-to-day life. Few major issues that a person who is having visual impairments normally has to face, can be identified with a detail study [7]. There are issues with daily living activities and also there are other problems faced in communication, reading and writing. Evaluation of space and the displacement is the main problem area that a visually impaired person normally faces. The exact idea of the environment and its surroundings can't be acquired straightforwardly. Guidance is needed for them to obtain the necessary

information that they need. Also they have problems with the activities, which need extended maintenance of visual attention.

The main challenge the visually impaired people face is, not knowing about the surroundings and what kind of objects are there around them. For detecting the obstacles, which lies in the path, blind people normally use a stick while they walk. From that approach they can only detect that there are objects on the ground [9]. Beyond that identification, they are incapable of knowing what is there in front of them and the size of the object. For an example, they can identify there is an object in front of them but they do not know whether it is a chair or a table without thoroughly checking the object. And that is time consuming as well as they only receive the abstract idea of the object.

Daily living activities of such person has become hard compared with the normal people who have good vision. Also they do not have much idea about the environmental space that they currently moving in. If only they are in a familiar environment, they can navigate without facing many problems because they have the entire picture stored in their mind. Once they are relocated in an unknown environment, they need to learn significant information about the location before they are satisfied to navigate on their own. In indoor environments such as hotels, airports, hospitals, things become harder for the visually impaired person. Since those interior spaces are large and have various kinds of signs, the visually impaired normally face more issues.

For obtaining additional information about better interpretation of the environment, computer vision is being focused and used [9]. Among various technologies used in assisting visually impaired people, the majority is targeted on mobility and obstacle detection [7]. Even though different developing technologies are being used for the attempt in helping visually impaired people, real time object identification is considered as a major difficulty [8]. Most of the related works are targeted on accepting inputs from the user surroundings and extracting information about the entities in that environment and provide the outcome by using auditory or tactile outcome [7].

2.2 Mobile access technology

Since the appearance of mobile devices, human computer interaction has changed significantly. Touch screens provide a great flexibility and a direct access to controls and information [12]. Computer assisted instructions have become a major future trend and it is being used in various mobile applications and software available in training, teaching and learning processes [13]. Mobile devices with cameras are ubiquitous and along with the modern computer vision techniques, new solutions to the problems discussed, can be identified [11].

Most of the assistive devices are financially out of reach for the population of visually impaired [15]. There was a need for some other approach, which is cost effective and can be used with common technologies. Also the people should be able to easily carry the devices whenever the places they visit. Rather than providing a solution, which needs complex hardware accessibilities, mobile device can be used to detect and identify the objects in the surroundings.

For example, if the user is in a super market, he should be able to identify the objects on the shelves in order to decide which one to purchase. By providing a mobile-based solution to the visually impaired people, this task becomes much easy. Since the smart phones have become essential to the people and it has become more close to a person's life, without any concerns they can be used as an assist in providing a solution to the visual impairment.

Even though mobile phones are nowadays popular, also among blind and visually impaired users, it is not easy to access those mobile applications like normal people. There should be extra adaptive interfaces and tools designed and developed for the people with visual impairments [13]. Many areas need to be explored in providing a mobile-based solution, which addresses the limitations that the visually impaired people currently face.

2.3 Related work

As object recognition has been subject to an extensive research area, there exists a range of solutions implemented in order to achieve the target of identifying and recognizing the objects. There is a system developed with the approach of time efficient cascade by using greedy algorithm to minimize the criterion and text detection algorithms along with that to provide a system, which can be used to help the visually impaired.

Even though there are object recognition solutions introduced, there is limited number of mobile-based approaches that provides the accurate features extraction of the detected objects. And also most of them are focused on providing the details of a captured image where the user has to point the camera towards the object without actually seeing it. That approach may work on some scenarios but can be failed most of the time.

2.3.1 Existing Mobile based Solutions

There are few similar applications that can be found in Appstore and Google Play where the needs of visually impaired people are being targeted and addressed. Following provides more information about those applications one by one.

LookTell is a commercial application for iPhone users which is supposed to recognize the object within the camera field of view that was previously stored in local database of objects' images. This application mainly focused on identifying household objects. The object templates, which are stored in the local database has to be captured by a sighted person in a predefined environment. The visually impaired person can point the mobile towards the object and recognize those household objects [14] [21].

There is an iOS application called iDentify in Appstore where user can point the camera towards an object and take a picture of it or load an image from the gallery. After that, it will provide the description of the previously captured or loaded image. It has stated that with the use of artificial intelligence, the description of the object is

being analysed and spoken out loudly. Internet connection needed to be available in order to acquire the result. Without Internet, this application cannot be used. [22]

EyeSpy is another similar application like the above mentioned mobile applications where the user needs to point the camera towards an object and then capture the image or obtain the previously captured image from the gallery. Then the image will be scanned and provide the user with information of what kind of image is that and where he can buy something similar to the captured image and provide more information about the related items and compare them. This app needs to have a internet connection to provide the results of the captured image [23].

CamFind is also a mobile visual search engine, which captures an image via the camera of the mobile phone, then provides the user a description about what is being captured and similar results of it. It has been a helpful app for the visually impaired people to identify the clothes and food items. Audio support is given in the app only after the user captures the image. This sends the captured image to a server and then only identifies what is being captured. It is a must to have a good Internet connection to use this app and obtain real time end results [24].

TapTapSee iOS application that is available in Appstore and the similar android application have also focused on the object recognition similar to the above apps. The user needs to take a picture of the image or load the image from the gallery and it will provide an audio description about the recognized object. It uses crowdsourcing to describe the picture captured by the users [5]. For that CloudSight.ai image recognition API has been used. This also needs Internet connection to identify the captured objects. This provides much better and detailed description about the captured image within 10-15 seconds. Compared with the other applications, TapTapSee can be mentioned as a better approach [25].

There are some few other similar applications where the object recognition is done using the help of communication with a server. Most of the above applications do not provide accurate information about the captured object. They mainly target on finding similar objects in the Internet where the user can purchase online. Those applications mainly focus on not only the visually impaired people but also other

people who have good vision. And because of that object identification have been achieved by sending the captured image to a server. To use those above applications, having a good Internet connection is a must where the images are being sent to the backend via an API and being processed there.

Following table 2.1 contains a summarized comparison among the features available in proposed system and the prevailing systems.

Feature	Proposed System	LookTell	iDentify	EyeSpy	CamFind	TapTapSee
Object Detection	Yes	Yes	Yes	Yes	Yes	Yes
Capture image of the object	Yes	Yes	Yes	Yes	Yes	Yes
Real time object Identification	Yes	No	No	No	No	No
Offline object identification	Yes	No	No	No	No	No

Table 2.1: Comparison of proposed system with prevailing systems

2.3.2 Deep Learning approach

Deep learning methods have recently influenced several application domains and significant performance improvement can be identified in those domain areas [18]. Many areas are focused by deep learning frameworks including computer vision, speech recognition, natural language processing and etc. Different deep learning frameworks optimize different features and different running performance.

The list of available frameworks includes Neon, Theano, Torch, Caffe, PyLearn, deepmat, CNTK, TensorFlow and etc [18]. They differ from one another with the aspects of training and deployment of deep learning algorithms. For instance, Theano provides automatic differentiation capabilities, which facilitates flexibility in modifying architecture for research and development. In Caffe framework, standard layers can be easily configured without hard coding which emphasizes ease of use [18]. Each framework has its own advantages and limitations.

TensorFlow is developed by Google, which is an interface for expressing machine learning algorithms, and an implementation for executing such algorithms. It is a C++ based deep learning framework along with python APIs developed and open sourced under an open source Apache 2.0 License by Google recently [16]. It is stated as a flexible system where it can be used to express wide variety of algorithms including deep neural network models. It supports deep learning to be used in mobile development. TensorFlow is designed for remarkable flexibility, portability, and high efficiency of equipped hardware. Many research are being started recently about this approach since it was introduced few years back. There is only few research papers published on TensorFlow since most of the research are still being carried on.

TensorFlow uses data flow graphs for performing numerical computations where the nodes represent mathematical operations and the edges represent multidimensional data array communicated between them [18]. TensorFlow has a flexible architecture that supports multiple backend, CPU or GPU on desktop, server or mobile platforms. TensorFlow also offers users the capability to run each node on a different computational device making it highly flexible. With better CPU scalability than other deep learning approaches, TensorFlow can be used for object identification purpose because it supports both large-scale training and inference [17]. Automatic differentiation and parameter sharing capabilities of TensorFlow framework has improved the framework to allow a wide range of architectures to be easily defined and executed [18].

Unlike the distributed TensorFlow framework, the other deep learning frameworks maps the computation onto a single machine. TensorFlow allows computations to be spread out across many computational devices across many

machines, and allows users to specify machine-learning models using relatively high-level descriptions. TensorFlow uses a hybrid dataflow model that borrows elements from each of these systems [16].

Following figure 2.1 shows the comparison of properties of few deep learning frameworks such as Caffe, Neon, TensorFlow, Theano and Torch.

Property	Caffe	Neon	TensorFlow	Theano	Torch
Core	C++	Python	C++	Python	Lua
CPU	✓	✓	✓	✓	✓
Multi-threaded CPU	✓Blas	x Only data loader	✓Eigen	✓Blas, conv2D, limited OpenMP	✓**Widely used**
GPU	✓	✓customized Nvidia backend	✓	✓	✓
Multi-GPU	✓(only data parallel)	✓	✓**Most flexible**	x Experimental version available	✓
Nvidia cuDNN	✓	x	✓	✓	✓
Quick deploy. on standard models	✓**Easiest**	✓	✓	x Via secondary libraries	✓
Auto. gradient computation	✓	✓Supports Op-Tree	✓	✓**Most flexible** (also over loops)	✓

Figure 2.1: Properties of several Deep-learning frameworks [18]

TensorFlow has a fast growing community of users and contributors making it an important deep learning framework within the community. Its high level APIs make the deep learning possible for almost everyone. The accessibility and the ease of use of the framework have gotten the attention of the developers compared to other deep learning frameworks.

2.4 Summary

Visually impaired people face diverse kinds of life challenges that normally sighted people take for granted. Computer vision based techniques can be used in providing an assistant for these people. Even though there are some existing applications, which have addressed the problems faced by, visually impaired in object recognition, they are not much accurate and need server support in providing the result. With the benefit of TensorFlowlibrary, we can provide a better outcome for those people who are suffering from visual impairments.

Chapter 03

Analysis and Design

3.1 Introduction

In any project, analyzing and design are the core aspects. That is the main idea, which is being generated throughout the previous chapters. This chapter contains information about the project analyzing and the design. Development methodology is further discussed in order to provide a clear idea about the proposed system. The analyzing part of this chapter contains the details about the system as mobile application and the justifications. The Design section explains the architecture of the system and a brief explanation of the workflow.

3.2 System Analysis

A massive area is covered in object identification. It can be categorized as indoor object identification and outdoor object identification. In each area that is classified, there is another set of large data to be considered. It can be complexed for real time object detection and identification in such a vast area of objects. For object identification, all the gathered data should be analyzed and then go through an identification process. Since there are many data to be analyzed, it is necessary to limit the scope and narrow down to a limited area.

Scope is limited to indoor object identification for this system and only considered the main indoor landmarks. Without considering all the objects in the first phase of the system, a better solution can be given by trying to identify a particular object in detail. Since the visually impaired need the support to identify special and commonly used areas, landmarks identification can also be focused in this project.

The system is all about improving the quality of life of people in the world who are visually impaired. The research has been carried out to implement an accurate object identification system with the use of deep learning approach. A mobile-based solution is being introduced as a solution to these people. The main

reason behind it is that the mobile phones can be carried to any place. There will not be any hazard in carrying a mobile device to the places they travel. Also in these days normally most of the people use a mobile phone in their day-to-day work.

3.2.2 Requirement Analysis

The main objective of this system is to provide a mechanism to assist in real time object identification for visually impaired people in an indoor environment. Providing a better guidance about the surroundings is the targeted outcome of this solution. By achieving that, it will improve the quality of life of people who are visually impaired. People who are blind can have a better idea about what is unseen for their whole life. Their burdens can be reduced since they do not need to have a support of another person in identifying objects and landmarks in indoor navigation. Also for the people whose vision is low or having troubles in the vision due to some injury, this application can be helpful in providing the necessary details of blurred out environment.

3.2.2.1 Non Functional Requirements

As the basic requirement of this project, a device is needed, which can be carried with the users without any burden. Mobile phones have become part of humans' lives. It should be easily carried around and the visually impaired people will not be unwantedly highlighted among the other people.

Accuracy of the system is considered as a main requirement of the project. People who are visually impaired depend on the output that is being provided by the system. Because of their inability to see the world, their lives are being affected by the proposed object identification system. Security is another main aspect that needs to be achieved from the proposed system. If the output of the system is wrong, their lives are directly affected. Because of that, such situations needed to be avoid and accuracy of the system needed to be increased.

Efficiency of the system is another main requirement of the project. The output of the system needed to be provided to the user real time. Lateness of the system will not be good since the user has to wait more time to get the result from the system. The user might use the application while they travel. Real time object

detection and identification is needed in such situations where the user gets the opportunity to get the details about the objects in the environment instantly.

The system should be more reliable because the visually impaired people are using the mobile application. They depend on the application in identifying the objects and the correct result should be transmitted to them when they use the application. The system should be tested perfectly before the application is released to the app store.

The system should have the ability to evolve with the time. Once the changes that need to be done are identified while the user use the system, they can be applied and provided to the users as a new release of the application.

3.2.2.2 Functional Requirements

The visually impaired people should not feel desperate because of their inability to see the world. As a solution to that, the camera of mobile phone is going to be used in identifying and detecting objects in the indoor environment. Object detection needs to be identified and then we need to recognize the detected object. For an example, if a fire exit sign is detected, we need to identify what is visualized from that detected object. Then feature extraction of that identified object needs to be carried out such that the visually impaired person acquire a better idea of what is it meant to be.

It is necessary to have a mechanism, which will be used to accomplish object detection and identification. A deep learning technique is considered in achieving this. Deep learning methods have resulted in significant improvement in performance in various application domains. Several software frameworks have been implemented to facilitate the development with machine learning [18]. From them, TensorFlow framework is a multi- layer neural network, which is composed of several computational layers. It process data in hierarchical fashion [19]. It is identified as a good mechanism that operated in large scale and in heterogeneous environment [17]. A need of such framework is highly essential in providing the proposed solution.

Since the visually impaired people cannot get an idea about the object identification and feature extraction without a guidance from an audio based solution,

all the information extracted need to be transferred in a proper way. For that, a text to speech conversion would be needed if the extracted information is not in readable format to these people.

3.2.3 TensorFlow as Deep Learning Framework

TensorFlow is an open source software library that is used for numerical computation using data flow graphs [20]. It acts as a supportive framework in helping the developers make lean mobile applications. It reduces code footprint and support in quantization and lower arithmetic that reduces the model size [20].

It uses a single dataflow graph, which maps the nodes of the dataflow graph across many machines [17]. The nodes in the graph represent mathematical operations, which is a unit of local computation. Each node has zero or more inputs as well as zero or more outputs [16]. The edges of the graph represent tensors identified as multidimensional data arrays that do the communication [20]. The dataflow graph of TensorFlow including the individual mathematical operations and their update rules with the input pre-processing is described in Figure 3.1.

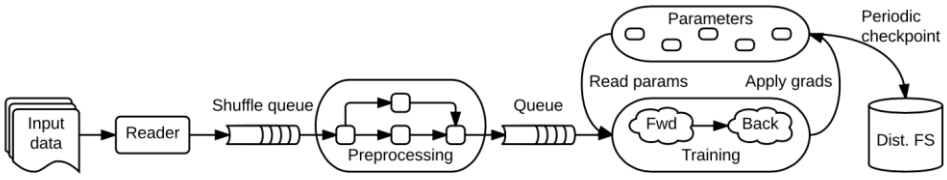

Figure 3.1: TensorFlow dataflow graph [17]

It is identified that TensorFlow differs from other dataflow systems in two aspects. The model supports multiple concurrent executions on overall sub graphs of the overall graph. And also the vertices of the graph may have mutable state that will be helpful in sharing between different executions of the graph without focusing on single execution [17].

Following figure 3.2 shows a feed-forward neural network, whichhas two hidden layers. The left panel shows a diagram of the neural network. Right panel

depicts the corresponding TensorFlow computational graph, which is graphically represented by the TensorBoard tool [19].

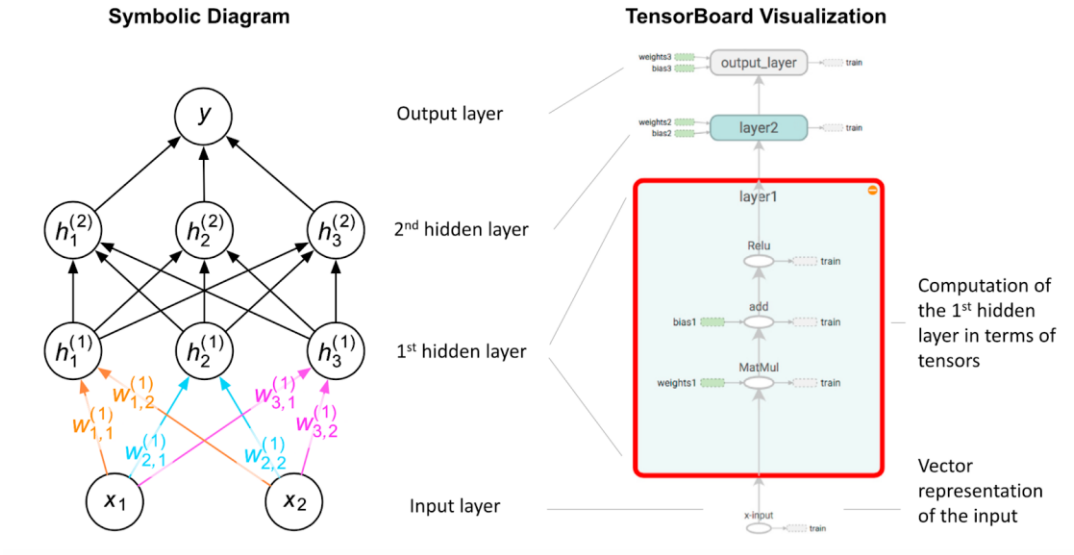

Figure 3.2: Feed-Forward Neural Network [19]

3.2.4 Mobile Machine Learning

Traditionally, deep learning has been associated with data centers and giant clusters of high-powered GPU machines. It can be very expensive and time-consuming to send all of the data a device has access to across a network connection. Running on mobile makes it possible to deliver very interactive applications in a way that's not possible when you have to wait for a response from network.

On-device deep learning can be categorized as follows.

- Speech Recognition - Audio based commands are given to the system with a speech-driven interface. Streaming audio continuously to a remote server to identify the provided audio is a waste of bandwidth. To reduce the time and save the bandwidth issues, small neural network can be built inside the device, which listens out a particular keyword. Once the keyword has been spotted, if it needs more computing power, the rest of the conversation can be transmitted over to the server in order to provide further processing. It reduces the background noises and the waste of bandwidth.

- Image Recognition - Camera applications are very popular among people because it’s very useful to recognize the content of the captured image or the camera image. To apply appropriate filters or label the photos, the help from image recognition can be used since they are easily findable after the process. Differently pre-trained models can be applied to recognize the objects inside the images by running them in mobile devices. Very fast and lightweight training can be used to reach it recognize specific objects with the pre-trained models.

- Object Localization - To detect the objects in an image and identify them, a pre-trained model of TensorFlow can be used. Drawing bounding boxes around the detected people in an images and track them over time can be achieved with TensorFlow. To detect a new object enters or leaves the scene, tracking plays a vital role and it can provide us the count of the objects that are present over time.

- Gesture Recognition - Either from images or through analyzed accelerometer sensor data, the applications can be controlled with recognized hand or other gestures. TensorFlow provides an effective way of deploying the approach by creating the models.

- Optical Character Recognition - Multiple steps are presented to recognize the text in images. First the areas where the text is presented need to be identified and then the letters needed to be interpreted in order to guess the words using a language model.To identify the letters, which are presented, segmenting can be used to divide the letters into individual letters and then a simple neural network can be applied to the bounding box of each.

- Translation - To translate from one language to another accurately and efficiently without an Internet connection, TensorFlow framework can be used. Descriptions about many different models in the literature can be found easily. Deep networks are very effective at tasks like this. Single graph to do the whole translation can be applied without using separate parsing stages.

- Text Classification - Our own categories or labels can be defined and train the examples. To suggest relevant prompts to users while they are typing or reading by understanding the meaning of the text, text classification can be used as a better option.

- Voice Synthesis - For providing users feedback or aiding accessibility, deep learning can be used to give the user a synthesized voice, which is similar to natural sounding speech.

3.3 Design of the system

In the design of the proposed system, production standards are mainly focused since it is designed for visually impaired people. It is a system, which evolves day-by-day since it is efficiency and effectiveness depends on the amount of objects that can be recognized through the system. The end user can see the system as a mobile application, but the complicated deep learning model is to be handled for the better outcome of the system. Sample data has to be feed to the model considering the objects in indoor environments and it has to be directed through an advance testing process since the information needs to be correct and effective. The basic design approach of the system is explained in figure 3.3

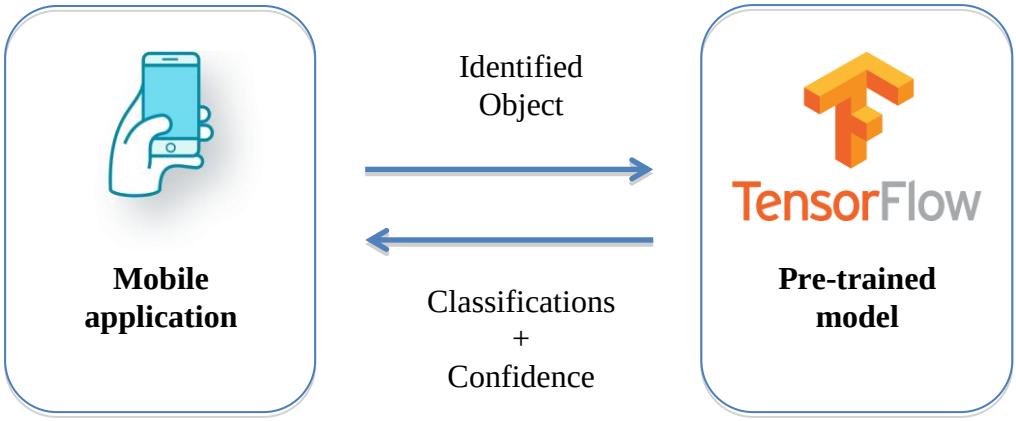

Figure 3.3: Basic approach of the proposed system

Three main areas are being focused in designing and developing the front end of the mobile application and the deep learning aspect of the system.

- Make it user friendly for the visually impaired people since they cannot handle mobile phones as normal people.
- An audio based guidelines need to be given since the users cannot visualize the object on their own.
- Making the system stable and reliable in terms of performance.

When considering the first design goal, standard human computer interaction principles are being followed and the patterns and methods used in existing mobile application are being referred as we can obtain a better idea about how the system should be enhanced by considering the drawbacks of the existing systems. Main focus of the application was to reduce the time spent by the visually impaired people in learning about the system. Even the advance features should be simplified and given to the users.

Second goal is to provide a better user experience for the visually impaired people with audio based guidelines in handling the application and providing the output of the identified objects.

Third aspect of the design goal is to make the system reliable and stable. For that TensorFlow framework plays a major role. It reduces the time spent in identifying the objects with the approach they have taken in computation.

Figure 3.4 describes the top-level system architecture of the proposed system and the connectivity between the end user, mobile application and pre trained models.

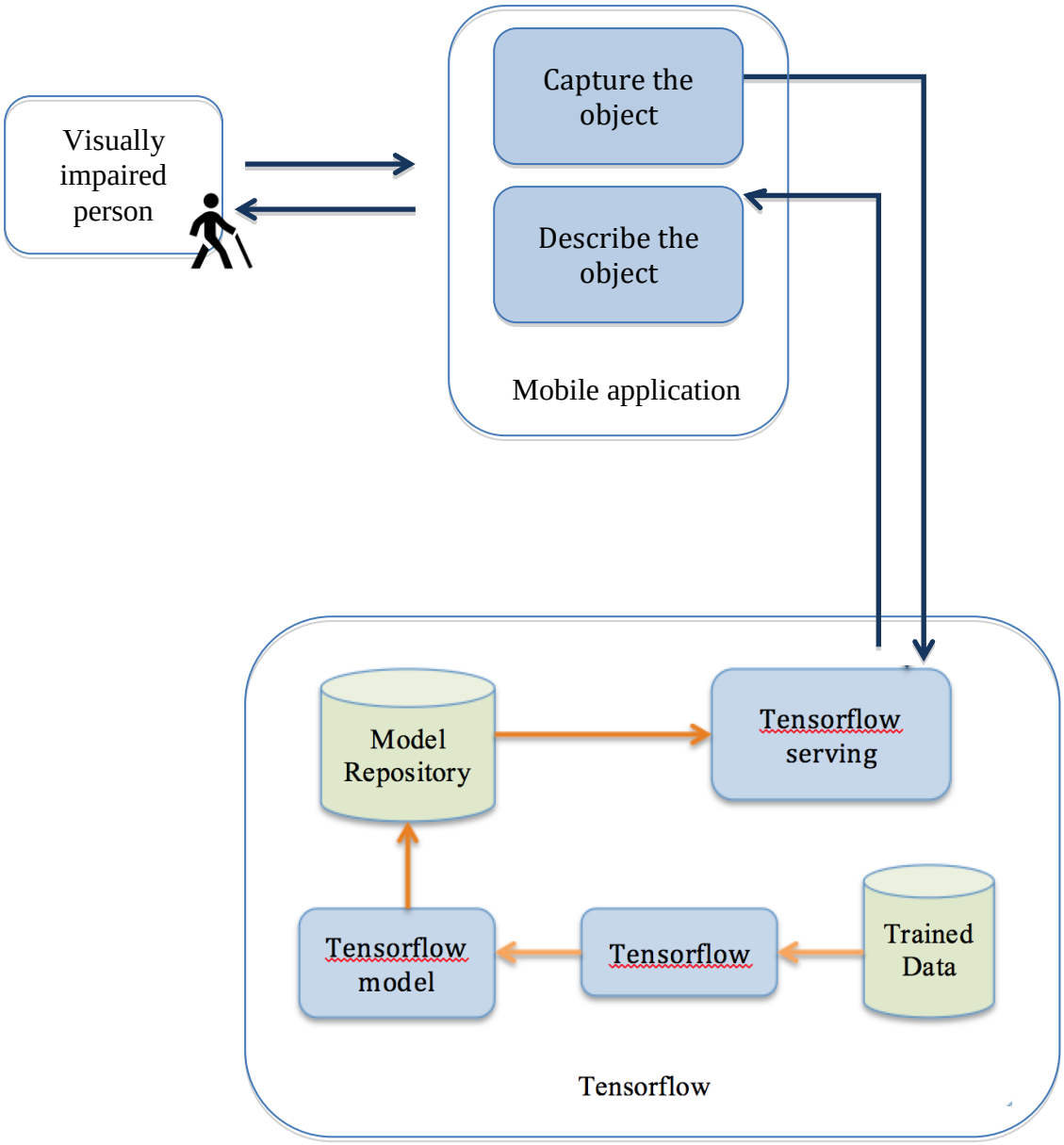

Figure 3.4: Top level system architecture

3.4 Summary

This chapter describes about the analyzing and design of the project. The system is a mobile application and it provides object recognition with the use of TensorFlow framework. The design of the system is highly considered since the system is designed for visually impaired people.

Chapter 04

Implementation

4.1 Introduction

This chapter describes about the developing tools and techniques used in implementing this system and the details of the implementation part of this system. The implementation can be categorized into two main sections as the implementation of the mobile phase and deep learning phase. Following section demonstrates the implementation details of each section of this system with their processes after discussing the tools and techniques adapted.

4.2 Developing Tools and Techniques

Intend of this project is to develop an iOS based application targeting the needs of visually impaired people which provides them the guidance in detecting and identifying the objects in indoor environment. In order to achieve this purpose a collection of the latest technologies is being used. Following is the list of technologies used in this project.

- Mobile (iOS) – Xcode (9.1) - Objective C, C/C++, CocoaPods
- TensorFlow deep learning technique - TensorFlow Experimental Pod

For the development of the mobile-based application, native development is targeted instead of hybrid applications because of the deep learning approach. The proposed solution needs to be efficient and reliable. Since the visually impaired people are going to use the application, the urge to provide a more user-friendly application becomes a major objective of the system.

Within iOS and Android platforms, iOS is targeted because of the encouragement of providing accessibility software in all most all the products. And also it comes with a simple interface and can be easily used by the visually impaired. iOS based applications are considered to be limited when considered with Android based ones.

Requirement of a free application, which will become a bridge to fill the gap between the need and the existing solutions, is achieved by the proposed approach.

4.3 Implementation of the System

Implementation of the system can be categorized into two main sections.

- Implementation of TensorFlow Support
- Implementation of the Mobile Section

4.3.1 Implementation of TensorFlow Support

TensorFlow is support is one of the main components of the developed system. There are few step that needs to be taken while integrating the deep learning network with the proposed mobile application.

1. Determine whether your problem is solvable by mobile machine learning -

   First a plan should be created to solve the problem idea and for building the solution. The problem should be actually solvable and it has to be clarified at the first stages of the problem solving stages. If the problem is not solvable with the proposed solution, a redesigning need to be done since training the computer will lead to difficult situation.

2. Create a labeled dataset to define your problem -

   A labeled dataset need to be created and define the problem that needs to be solved. This step is considered as more important than selecting the appropriate model. The dataset needs to be a representative of the actual use case. Labeled dataset need to be efficient and accurate as possible and to achieve that appropriate tools can be used. After initial labeling, errors and difficulties can be identified and that experience can be used to change the labeling and data capture process to avoid the unwanted problems. After consistent labels are being created, the external raters can be guided in running the same process.

3. Pick an effective model for the problem -

The final step is to pick up an effective model to be used. If any other person is training a similar model, training from scratch can be avoided and it will be a great help. TensorFlow has a repository of models that are being implemented and it can be referred while picking an effective model. A simple model can be picked once the labeled data set is granted and then start the process. It is said that the best results can be achieved when the iterations can be done quickly. If shorter time is taken to train a model and run it in a real application, the overall results can be lead to a success. The results can generate a failure if there's a mismatch between the dataset and the real usage. The accuracy of the system gets decreased with a situation like this. A consistent user experience can be created by prototyping the end-to-end usage as soon as possible.

4.3.2 Implementation of the Mobile Section

Another main component of the project is the mobile section that interacts with the targeted visually impaired people in helping them with the object identification. It acts as a bridge between the people who are visually impaired and the surroundings in helping with their needs.

For the implementation of this component, objective C was used as the core technology with the use of C/C++ components. Usage of Cocoa pods has been used to access the TensorFlow Experimental pod.

The most necessary aspect that has to be considered in developing this project is that this project is not going to be developed for the normal people with good vision. The targeted segment of this application is the people who are visually impaired. The mobile section is done in a way that facilitates the need of those people with audio guidance throughout the system.

Human interaction with the application needed to be revised after applying a certain approach since the application is being developed considering the people with low vision or blind. Accessibility feature is available in iPhones to support the people who are differently abled. People who are visually impaired can use the accessibility guidance to open an application and know the content of that. By using the voice assistant feature "Siri" provided by Apple, a person can provide commands to open an

application on behalf of the user. Once the command is given as “Open (Application name)”, it will direct the user to the application by opening it in the next second.

When the application is loaded, camera view is given to the users to detect the objects directly. It captures the objects that are being visualized in the camera view and then identify them with the trained model stored in the application itself. The main benefit of the application is that it supports for the offline mode. The captured images are not sent to a remote server and then identified the objects like other application has been done. The objects that are being visible through the camera view of the application, are mapped with the trained dataset where one thousand data are being labeled to recognize the objects detected.

The first image in Figure 4.1 shows the real time recognition results when the user detects the object without freezing the frame. Second one shows the result for same image that comes after freezing the frame. User can freeze the live camera view and obtain the real time recognition results.

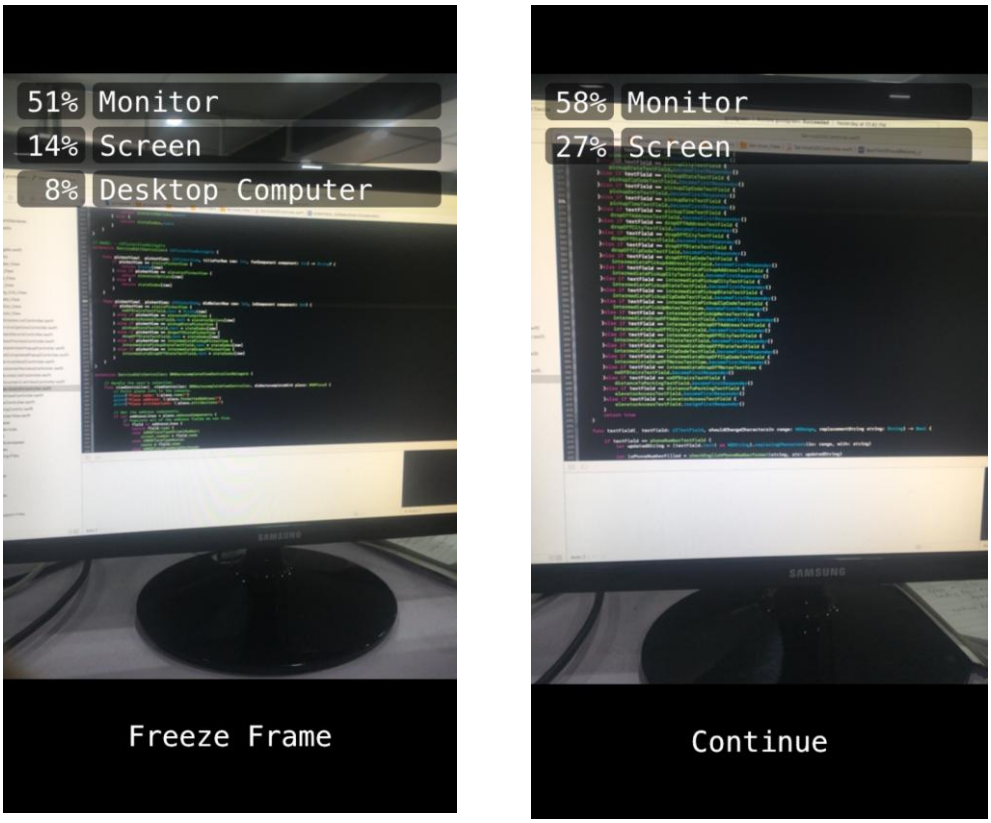

Figure 4.1: Real time recognition of same object before freezing the frame and after freezing the frame

Figure 4.2 shows the results that are visible for several objects that are being detected through the system. The results that are being identified are provided to the user via audio output.

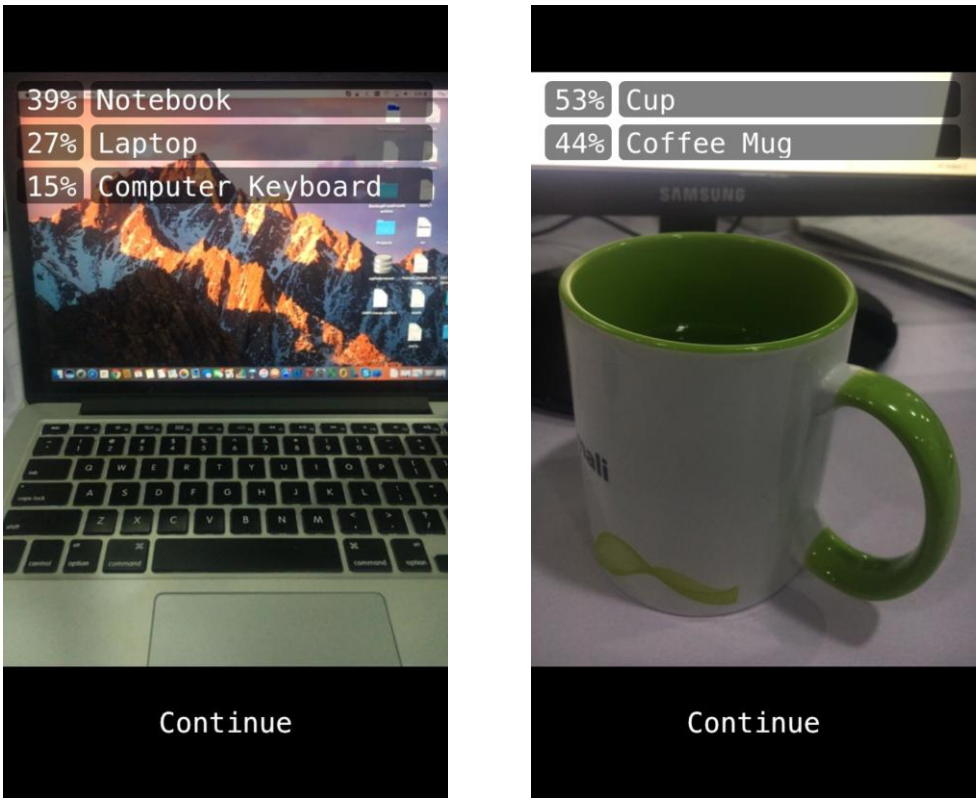

Figure 4.2: Real time recognition of different objects after freezing the frame

4.4 Summary

This chapter describes about the implementation of the project. It mainly focuses on the mobile section of the application. Real time object recognition is achieved through the system using a live camera and provides the results without connecting to a remote server. All the object identifications are done in offline mode. The implementation of the system is highly focused on proving a better approach for the visually impaired people.

Chapter 05

Evaluation and Results

5.1 Introduction

TensorFlow based object identification for visually impaired is proposed as a mobile application which facilitates in object detection and identification in an indoor environment. It provides a better guidance to the people who are visually impaired in detecting objects and then identifying them on behalf of those people. Main goal of this project is to provide a better life for the people who are not privileged to see the world as we normally see. This chapter is dedicated to discuss on the evaluation of project with respect to the objectives as well as the data that is being used as assistance to the project.

5.2 Evaluation at Different Levels

The main purpose of the project is to provide a better guidance in object identification for the people who are visually impaired. To achieve the goal, TensorFlow deep learning framework is being selected since it is more feasible compared with other known frameworks. Since the project is targeted for the people who are visually impaired, the object detection and identification needs to be accurately given. Project was monitored on a regular basis during the course of the project life cycle, with the help of Gantt chart as well as meetings with supervisor. In the middle of the project life cycle period, using identified needs system was clearly defined as well as designed.

The evaluation of the solution has been carefully planned and categorically assessed. It was done with the help of an evaluation plan. Evaluation plan address all the stages of the development life cycle of the project. All necessary activities were broken down and the primary decisions were taken place considering the project

activities. At very first stage by submitting written project proposal, requirements feasibility and clearness of problem domain, appropriateness of technologies and adequate resources were identified.

The main evaluation of the system begins at the first phase of the project with the start of gathering data that is related to the domain. At the initial stage, indoor objects were targeted as the test data so that the evaluation of the system is focused on considering only that particular targeted sub domain. Since there can be huge amount of objects to be considered, the domain was reduced to a targeted set of data where the identification can be easily managed. TensorFlow framework is providing a set of test data that is being used as the example models. For the initial implementation of the project, those test data has been used since we can provide a better outcome with the given thousand sets of data.

When developing the application side of the project, we mainly focused on the interaction side of it with the users since the people who are visually impaired are the mainly targeted users of the system. The application needs to be compatible to be used by those people and audio navigation can be given when they use the app. When developing it, many facts had to be considered such as how they can interact with it and how the output needs to be given.

Object identification is the most important section of the app considered with the application side. Since the objects needed to be identified accurately and given the output to the users in an insignificant amount of time, the implementation of the object identification is prioritized in this project. When using most of the frameworks in object identification, they grab an image from the mobile side and then send it to a server to recognize it. It takes significant amount of time and the user needs to always use an Internet connection to obtain the results. If the network connection is slow, the output will come to the user after some amount of time and the user will have to wait few minutes to know exactly what object is there around them. As a solution for that, with the help of the framework that has been used, real time object identification was achieved. To provide the output more efficiently and effectively, the data set has to be improved by adding other objects to the model.

5.3 Sampling

In order to collect data for the evaluation process a representative sample has to be identified. This sample was focused on indoor object identification and only the highly used objects are targeted in creating the models in the first phase of implementation. By stratifying the collected objects into similar categories, sample groups are being formed. Since TensorFlow framework is providing a sample indoor dataset, it is being used for the modeling. As the initial stage of the system, only thousand sets of data is being used. While evaluating the whole system, it is essential to consider only the objects that are being used as the data set in modeling the TensorFlow graph.

5.4 Types of Evaluation

There are different types of evaluations methods depending on how the system is evaluated and the purpose of evaluation. A user-based evaluation is considered in evaluating the system. And to accomplish that, the progress of the identification of the objects is being continuously monitored. To measure the performance of the network, the evaluation of the model can be done for both the training and test datasets.

- What is evaluated?

How accurate an object is identified when there are several set of objects in the captured frame? How the identification is improved with the proper approach and how the quality of life of the people who are visually impaired has improved with the use of the proposed system?

- What is the purpose of the evaluation?

The purpose of the evaluation is ultimately to achieve the project goal and the objectives. People who are visually impaired should be able to overcome the difficulties in identifying the objects. Therefore it should be evaluated whether the goals are achieved properly or not.

- Who is interested in the evaluation?

Interested parties of the evaluation process are basically the people who are visually impaired. If the project is successfully achieving the objectives and the goal, then this project can be extended to deliver.

- How will they use the findings?

People who are visually impaired can use the findings of the evaluation result in identifying the objects. How far an object is being identified and how accurate the results are.

- What questions should be answered?

    - For what extent the objects are being identified?
    - How much identificationis correctly given and how accurate the results are?
    - Up to what kind of features of an object can be extracted to the people who are unable to see them?
    - How far the real time object identification is achieved from this approach?
    - What kinds of enhancements should be applied and how the approach should be taken?

5.5 System Testing

Throughout this development lifecycle of this project, a series of system tests and validations were done in order to ensure the quality of final system. During the implementation of the system each modules were tested individually. After completing integration of modules, the whole system was tested as a whole.

As system tests mainly graphical user interface testing, error handling testing, and performance testing were done repeatedly. Outcome of one test was used as the base for the next testing phase. These testing helped to find out problems and bugs in the system and develop the system in a better manner. The developed application is

mainly focused on visually impaired people and all the system testing is done accordance to that scenario.

5.5.1 Tested Devices

The system can be installed into both iPhone and iPad devices and all the Apple devices have the ability to install the developed application. Implemented system has been tested in various Apple products.

Tested Device	Operating System
iPhone 5s	11.2.6 / 10.3.1
iPhone 6s	11.2.6
iPhone 6 plus	11.2.5
iPad 2	9.3.5
iPad Air	11.1

Table 5.1: Tested devices

5.6 User Acceptance Testing

User acceptance testing plays a major role in evaluation of a product. The user acceptance testing for this project was done mainly by providing a prototype of the product for few users and getting their feedbacks.

The implemented system targets a specific user category. It is a product for the people who are visually impaired. However the importance of this type of a system may not be understandable at once for a person with lesser technical knowledge. Therefore for this user acceptance test, group of people with some technological knowledge background about using mobile phones were selected. Since it's difficult to find a user base that are visually impaired and familiar with mobile phones, system was tested among both the people who are visually impaired as well has people who have good vision.

A questionnaire was prepared in order to evaluate the implemented system. Feedback from visually impaired people was recorded and filed on behalf of them. Following characteristics were targeted in the given questionnaire.

- Usability of the system - For how extent the visually impaired people can use the system and find it useful. The level of ease to learn and what kind of problems they face.
- Reliability of the system - The extent to which the system is expected to perform its intended functionality.
- Accuracy of the system - To which extent the response of the system is correct and identify the objects in the environment.
- Effectiveness of the system - Time taking to provide a result to the end users and how efficient and effective the system it. Less waiting time or had to wait more than they expected.
- Importance of the system to the society - The extent to which this system is useful to the society.
- Final evaluation of the system as a whole - Overall idea of users about the object identification mobile application.

There is an open question at the end of the questionnaire to collect general comments on the system from users. The feedback-collecting questionnaire is attached in the Appendix section under Appendix A.

Feedback collected using the above mentioned questionnaire is analyzed and finally summarized into a table. Result of this evaluation is shown in following chart Figure 5.1.

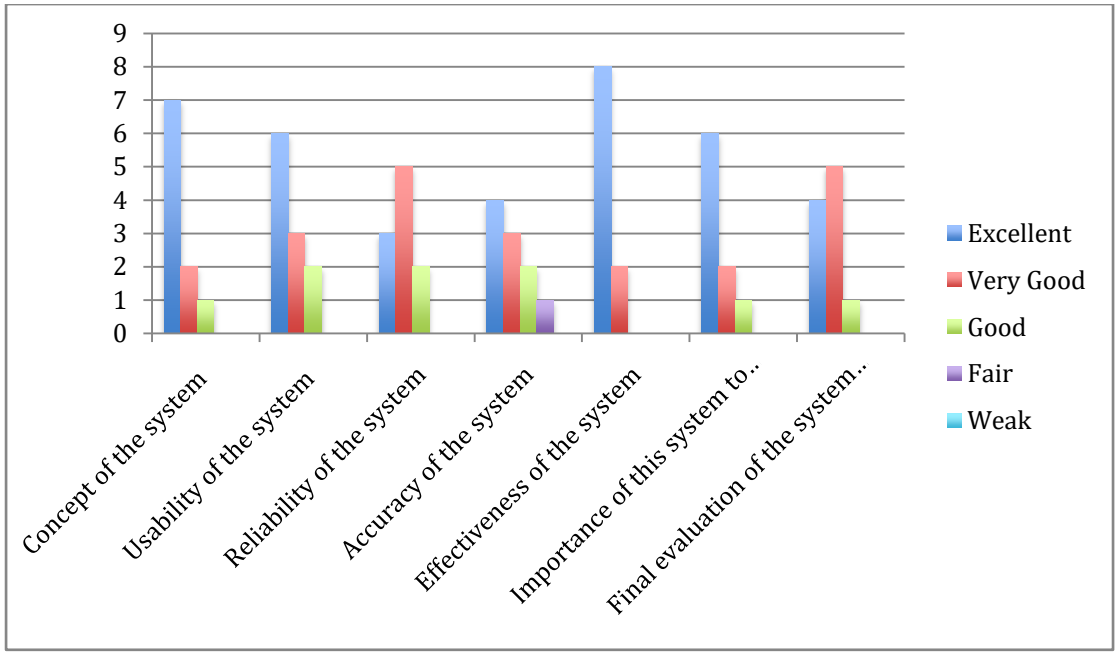

Figure 5.1: Evaluation Summary as chart

According to the results of evaluation it was possible to identify areas where further improvements are needed. Results show that the accuracy of the system and reliability has to be improved.

Also it is noticeable that many users have decided that the TensorFlow based object identification system for mobile devices are an important concept and such a system is important to the society.

5.7 Object detection evaluation

Along with the user acceptance testing, object detection needs to be evaluated. It is achieved by using different kinds of objects, which has similar features.The data that is being fed to the deep learning network needs to be identified even the shape and properties of the object iss different from each other. When we consider a water bottle, it can be in different shapes and different sizes. The color and other properties can be changed based on each and every manufacturer. The object needs to be identified by the system even the characteristics of that object is bit different with the data feed in the system.

For the object detection evaluation, certain set of objects has been used and the evaluation is done based on the accuracy of the system. Following are the objects that

we have focused on doing the evaluation and the amount of items from each that we have used.

- Water bottle – 20 items
- Pen – 15 items
- Keyboard – 6 items
- Laptop – 8 items
- Mug – 15 items
- Monitor – 5 items

The developed application is used to identify various kinds of objects and they are being monitored and evaluated. Along with the user acceptance testing, an evaluation based on the objects is necessary in measuring the accuracy of the system.

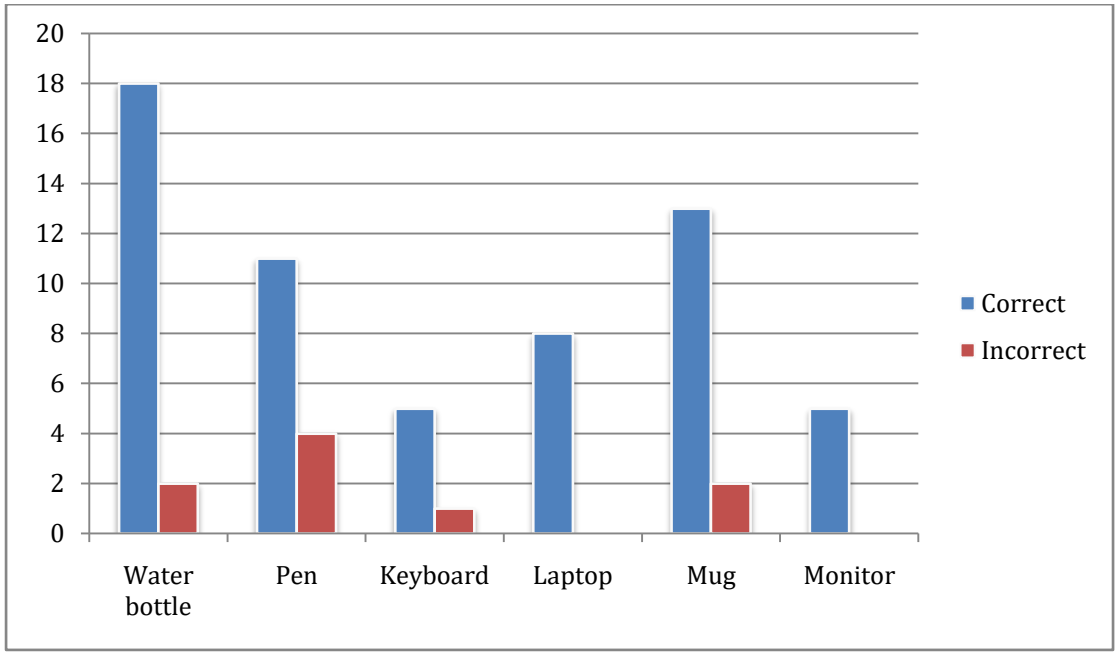

Figure 5.2: Object based evaluation summary

Evaluation results appeared to be more accurate for some selected objects. And some were identified as incorrect because of the difference of the properties of that particular object. The outcome of the evaluation emphasis that the research needs to be focused more about the diversity of the objects.

5.8 Summary

In this chapter all the evaluation criteria of the system along with the results gained by evaluations are presented. The summary of feedbacks from questionnaire is presented as a chart. Finally it describes areas to be improved based on the results.

Next chapter will be focused on conclusion of the system based on the achieved objectives, problems encountered and limitations of the system. Also further expandability of the system is discussed.

Chapter 06

Conclusion and Further Work

6.1 Introduction

The main goal of this project was to develop a mobile-based solution to visually impaired people in object detection and identification in an indoor environment. Various aspects of the project are being discussed throughout the whole dissertation. A full description about the problem and the proposed solution for that problem, the final product as well as the technologies being used are being discussed. The design of the project, its implementation, and the evaluation was the main focus in last few chapters.

This chapter mainly will conclude the dissertation and it is dedicated to provide a description about the overall achievement of the project and also the objectives achieved and problems encountered throughout the project under the scope and time and actions taken to overcome those problems. Finally how the project can be further developed in order to eliminate the problems encountered, the future extendibility is being discussed.

6.2 Achieving the Objectives of the Project

Accurate object detection and identification is essential for the people who are visually impaired. Those people are facing many problems in not having the opportunity to see the world or get any idea about the environment they live. Their lives are really hard because they have to live with many obstacles and problems. It is highly necessary to be aware of the problems that they are facing and provide a better solution and guidance. From this project, it tries to address those problems by providing a more accurate object detection and recognition technique.

The system ensures that the visually impaired people get the proper outcome and guidance. The identification process of a particular object via TensorFlow framework ensures them a quality life that they always seek. From this proposed

solution, the people who are either fully blind or partially blind can get a better idea about what they are not capable in seeing.

The aim of the project, which is mentioned in a previous chapter, is achieved successfully with the proposed solution. Following paragraphs further clarify how each objective of this project is successfully achieved.

One of the major objectives is to provide a mobile-based application, which fulfill object detection of surroundings in an indoor environment. A mobile application is developed with the help of TensorFlow framework and it provides the ability to detect and recognize the objects, which are in a particular environment. Users have to open the developed application and aim their phones to the direction that they wanted to get the description. Then the objects that are in that aimed direction are being identified by the system and let the users know the details of the surroundings.

An audio-based guidance is provided for the users to let them know what are the objects that are being captured by the application. Since the system is mainly targeted to the users with visual impairment, main objective of the project was to provide a simple mobile application, which can be easily used by them. If it's a complicated application with so many buttons and views, people who are visually impaired won't be able to even navigate from one view to another. What they expect from this application won't be fulfilled with a complex solution.

Essentially, a data set that's being provided by TensorFlow is used to develop the system. They have provided indoor object data set with 1000 objects. We have used that data set in feeding the system. With the given data set, we identify the objects that can be seen in the surroundings and provide an audio based description of the objects been detected. The people who are visually impaired can get a better idea of what they do in their life and can identify the objects that they need to use without any hazards that they were facing before.

Generally, the objectives of the project are accomplished as it was planned. Through this system people who are visually impaired get the opportunity to see the world and get a better idea of the surroundings. With all the achievements that are

being accomplished, it's possible to say this solution has successfully achieved all of its objectives.

6.3 Problems Encountered in the Project

When the project was flowing towards its success, various problems were occurred due to many reasons. They have to be addressed continuously in the process of the development of the system.

Gathering data about objects in the environment to create the data model was significantly hard. There a many objects that needs to be identified. To minimize the challenge that had to be faced, scope was reduced only to identify a particular set of indoor objects.

Another major problem occurred is the accuracy of the system. In this system, accurate information has to be provided as the outcome since the people with visually impaired depend on such data. That was one of the major challenges, which required immediate attention during the implementation of the system.

The TensorFlow framework that has been used in object identification provided an experimental Cocoa pod to be used in the implementation of the mobile applications. It was quite large since multiple platforms were bundled together in creating that.

6.4 Limitations of the Currently Developed Solution

The system is being implemented as an assist for the visually impaired to identify objects in the environment. It is being introduced to give a value to the lives of those people. When a visually impaired person faces an issue in identifying the objects, they can use the implemented system to guide them in order to discover what they need to know. The quality of life of these people can be improved with a solution like this.

Number of data that were being used to feed in to the model is one of the limitations of the system that is being developed. Only a data set with one thousand of objects is used in the implementation of the system. Main focus of the implementation

was to provide a better object identification and the amount of data that is being used is currently limited to that amount. Because of the evolving manner of the system, the accuracy of the system is proportional to the data that is used to build the model. The outcome of the developed system may not be reached up to the expectation level since some objects may be identified incorrectly because the system may not have trained for detecting that particular object.

The system is developed only targeting the iPhones and it's another limitation that the current system has. The system needs to be developed for other mobile platforms as well. Another main limitation of the system is that the ability of using the mobile phones is considerably small among the visually impaired people. With the emerging technology, this limitation can be changed in few years.

Even though there can be limitations of the currently developed system, they can be addressed and eliminated by taking necessary actions and further developing the system.

6.5 Further Extendibility of the System

As the further development of the system, we mainly focus on how to overcome the limitations that are being discussed in the previous section. There were two major limitations discussed in there and those have to be addressed one by one to provide a better system by further developing it.

The system is having a limitation with the amount of data being stored. Increasing the objects in the training data set can reduce the limitation. More indoor objects contained in the system can increase the accuracy of the system. Eventually the system will grow with more data and then the efficacy and the accuracy of the system will be increased.

Introducing the developed mobile application for other mobile platforms can reduce the second limitation of the system. It can increase the usability of the system among the people who are visually impaired. They might be familiar to another mobile platform and providing the system for other mobile operating systems, the usability can be increased with the time.

6.6 Summary

The conclusion of a successful project always marks the beginning of some more innovative further work. As this chapter declared, this system can be declared as a success, of which the main goal was to support the people with visually impaired in identifying the objects in an indoor environment. The resolving software has been developed to the expected standards to output the expected results by achieving the objectives frequently.

A particular project always carries problems when reaching its goals, challenges in implementation, limitations, and its further evolution. Even though some problems were encountered in the development phases, the project was lead to the final destination of illumination. However, the application of the concept that was addressed by the project is bit challengeable and the ultimate goal was achieved by the extreme effort. The currently developed system also comprises few limitations when implementing it in the real world. However, the further extensibilities of the system are being identified and those can be involved to provide better solutions in the future.

References

[1] “Visual impairment and blindness”, World Health Organization, Oct 2017

[2] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," British Journal Ophthalmology, 2011.

[3] E. Brady, M. R. Morris, Y. Zhong, S. White, and J. P. Bigham, “Visual challenges in the everyday lives of blind people,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 13*, 2013.

[4] A. L. Guerrero, F. Vasquez 2 and S. F. Ochoa, “An Indoor Navigation System for the Visually Impaired”, *Sensors*, vol. 12, pp. 8236–8258, 2012

[5] Q. Lin, R. Jiang, S. Qu, “Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio”

[6]A. Martinez-Sala, F. Losilla, J. Sánchez-Aarnoutse, and J. García-Haro, “Design, Implementation and Evaluation of an Indoor Navigation System for Visually Impaired People,” *Sensors*, vol. 15, no. 12, pp. 32168–32187, Dec. 2015.

[7]H. Jabnoun, F. Benzarti, and H. Amiri, “Object detection and identification for blind people in video scene,” *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015.

[8] K. P. Bhure, J. D. Dhande, “Object Detection Methodologies for Blind People”, *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 5, issue.1

[9] N. Mohajeri, R. Raste, S. Daneshvar, “An Obstacle Detection System for Blind People”, *Proceedings of the World Congress on Engineering 2011*, Vol. 2, July 2011

[10] S. R. Pampattiwar, A. Z. Chhangani, “Smartphone Accessibility Application for Visually Impaired”, *International Journal of Research in Advent Technology*, Vol. 2, No. 4, April 2014.

[11] K.Rector, S. Azenkot, R. E. Ladner, J. O. Wobbrock, “The Need for Research on Mobile Technologies for People with Low-Vision”, April 2013.

[12] J. S. Sierra, J, S, R. de Togores, “Designing Mobile Apps for Visually Impaired and Blind Users Using touch screen based mobile devices: iPhone/iPad”, *The Fifth International Conference on Advances in Computer-Human Interactions,* 2012.

[13] A. Aher, K. Musale, S. Pagar, S. Morwal, “Implementation of Smart Mobile App for Blind & Deaf Person Using Morse Code”, *International Journal of Research in Advent Technology,* Vol. 2, No. 2, Feb 2014.

[14] S. Parkhi1 , S. S. Lokhande, N.D.Thombare, “Vocal Vision Android Application for Visually Impaired Person”, *International Journal of Science, Engineering and Technology Research (IJSETR),* Vol. 5, Issue. 6, June 2016

[15] H.R. Arabnia, R. Jafri, S. A. Ali, “Computer Vision-based Object Recognition for the Visually Impaired Using Visual Tags”.

[16] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane´, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Vie´gas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng. “TensorFlow: Large-scale machine learning on heterogeneous systems”, 2015.

[17] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane´, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Vie´gas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng. “TensorFlow: A System for Large-Scale Machine Learning”, Nov 2016.

[18] L. Schott, M. Shah, N. Ramakrishnan, S. Bahrampour, “Comparative Study of Deep Learning Software Frameworks”, March 2016.

[19] A. Goldenberg, L. Rampasek, “TensorFlow: Biology’s Gateway to Deep Learning?”, Jan 2016.

[20]https://www.TensorFlow.org/

[21]https://itunes.apple.com/us/app/looktel-recognizer/id501088555?mt=8

[22]https://itunes.apple.com/ca/app/identifi-object-recognition/id1135223189

[23]https://itunes.apple.com/us/app/eyespy/id938516167?mt=8

[24]https://itunes.apple.com/us/app/camfind-visual-search-powered-by-cloudsight-ai/id595857716?mt=8

[25]https://itunes.apple.com/us/app/taptapsee-blind-visually-impaired-camera/id567635020?mt=8

APPENDIX A

Questionnaire used for evaluating user acceptance

Evaluating Object Identification Mobile Application

Section A: General Information

1) What is your age group?

Less than 20 years 20 – 30 years

30 – 40 years above 40 years

2) Gender

Male Female

3) Are you familiar with using mobile phones?

Yes No

4) How good is your eyesight?

Normal Vision Moderate vision impairment

Severe vision impairment Blind

Section B

	Criteria	Excellent	Very Good	Good	Fair	Weak
1	Concept of the system					
2	Usability of the system					
3	Reliability of the system					
4	Accuracy of the system					
5	Effectiveness of the system					
6	Importance of this system to the society					
7	Final evaluation of the system as a whole					
Comments						