Project Synopsis

Project Title:

"A Visual assistant chatbot for visually impaired"

Submitted by:

Pramukha R N
4NM16CS101
pramukhrn@outlook.com

Rahul D Shetty
4NM16CS111
35rahuldshetty@gmail.com

Shetty Yashas Shashidhar
4NM16CS137
yashasshetty111@gmail.com

Preferred Project Guides:

Name	Signature
1. Dr. Venugopala P S	
2. Dr. Roshan Fernandes	

Submitted on: **19th August 2019**

Abstract

A key challenge faced by people who are partially or completely blind is perception and navigation of the environment that they are not accustomed to. Travelling to an unfamiliar place or merely walking down a crowded street or can be a challenge. As a consequence, many people with impaired vision travel with a friendly person or a family member while navigating an alien environment. This proxy person helps them in their navigation, describes the external environment to them and thus helping them in their external cognition. We propose a chatbot assistant which plays the role of this person. Our "Visual assistant chatbot" acts as their artificial eye describing and summarizing the external environment in real-time.

The "Visual assistant chatbot" can be integrated into an android app which can be launched easily with a single click from the phone or can be launched through voice command. The person might ask summary or questions like *"What do I see in front of me?"*, *"What food is on the plate?"*, *"What are the people doing?"* and get answers in real-time. The chatbot also describes any changes in the current scene and also notifies the user about any warnings. This chatbot can be used for a wide variety of uses from reading the price of an item to helping in setting the time of a washing machine, describing the type of beverage in the cup to identifying what the giraffe in a zoo is eating. This chatbot leverages the power of computer vision and natural language processing to liberate any visual burned that visually impaired people may encounter and makes their everyday life easier.

Objectives

1. Design and implement a Visual Question Answering model which takes an image and a text question as input outputs an answer.
2. Create an interface for speech to text interconversion and prompts for the user.
3. Implement scene summarization to provide a quick summary of the current scene.
4. Implement scene change detection and warning indicators for potential dangers.

Future Development

1. Using smart glasses with a camera and running the model on the cloud for increased convenience.
2. Decreasing the speech to text and text to model output error rate and obtaining more specific answers.