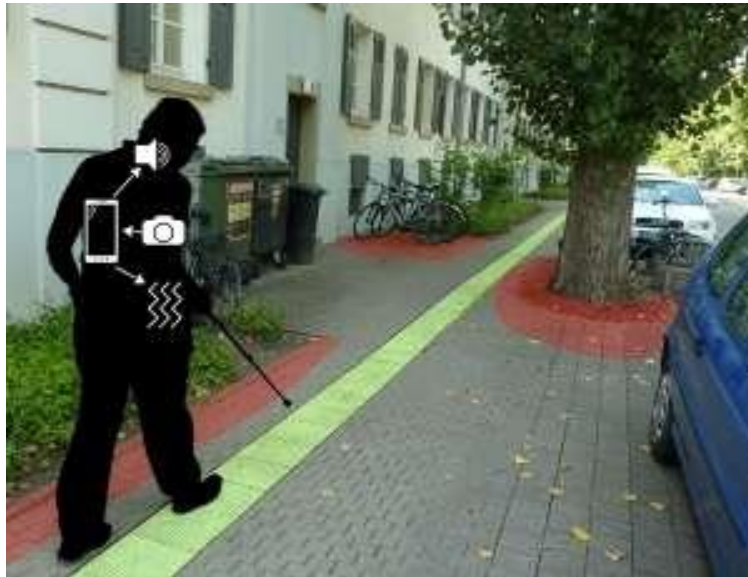


ENVISAGER: SIGHT FOR BLIND AND VISUALLY IMPAIRED

Project Phase - 1



Submitted on **27 February 2017**

By **Team 7 (Team Supreme)**

Nandanamudi, Jyothi kiran (27)

Gatiganti, Sai Venkatesh (9)

Gumma, Madhuri (13)

Nallamotu, Naveena (26)

I. Introduction:

In this document, we intend to provide our objective, significance and features of the project. We also discussed our goals and objectives for the application named “Envisager” and what motivated us for choosing this application. We aim to incorporate some added features to make this application notable.

The main objective of the project is to help the blind and visually impaired people crossing roads by detecting the object nearby and other obstacles by using the smartphone. The smartphone is used to capture the video of surroundings.

II. Project goal and objectives:

1. Significance:

In this hustle and bustle of modern life, even people with no disabilities find many things as a hindrance for completing their activities and daily routines. However, people who are visually disabled are facing many difficulties than the normal people without any disability. As we know, the social constructs are not always designed by keeping them people in mind. They do need support in crossing roads, detecting the object nearby and other obstacles. Even though they use canes, they cannot identify objects above their waists, which is why there is a need a smart way of identifying the object they come across.

With the same motivation, companies like Siemens started building application that would help visually impaired people to navigate through busy roads by considering the GPS and identifying the persons location and help them to reach their location. For the past, few years there are many inventions like smart canes that detect object impact and electronic glass (eSight) that lets the people see but, these items costs a lot and not all can afford to have one.

However, with an android app like ours, people with smartphone can have it and we could help most of them to perform their daily activities without running into any trouble. Our main goal is to develop an application which would allow the blind people to take the pictures through camera and then our system would detect the image captured and identify the object and give them the audio reply describing the object or the naming the object etc., Moreover, all the features can be used without spending a dime and all they required is to install the application.

Objectives:

- a.** To develop a user-friendly application.
- b.** To provide a smart application that would detect the object accurately.
- c.** To notify the user about the object with the voice that the person can understand.
- d.** To test the time taken by the process in different methods (using Clarifai API, Spark Machine Learning, Deep Learning) and choose the best method for our application.
- e.** To develop an application that can be operated with a minimal cost.
- f.** To develop a scalable project

2. Features: Use Case/Scenario

This project has two different feature (a) Image to text conversion (b) Text to speech conversion.

(a)	Image to Text Conversion	Used Clarifai API to convert the image into text in the form of annotations
(b)	Text to Speech Conversion	Android inbuilt text to speech conversion api is being used

Table 1: Features of the project phase-1

III. Approaches

In the phase- 1 we intend to use the Clarifai API to create annotation of image and by using text to speech conversion API, we intend to convert the created annotation into a speech.

We developed an Android Application by which an image can be captured and is passed to clarifai API. The image or the video captured while using this android app will serve as an input the response will serve as a partial output where we intend to filter the annotations and pass these to the text to speech conversion API which will read out the image captured.

Input: Image

Output: Audio (Image Explanaiton)

1. Analytical Tool

Clarifai API:

The Clarifai API opens up the whole new process of data analysis. Clarifai provides a variety of pre-trained models that we can use to make predictions. These models were thoroughly tested and been used by many data processing industries. And apart from these, In Clarifai we have chance to build or customize our own models. It classifies all the images and it has ability to handle the large amount of data.

The API accepts a video, image or an audio as an input and it will break down the media to analyse it and to use the information to configure and extract the words the media may contain. The API uses machine learning techniques which would improve with time and it gives the output to perform any further actions.

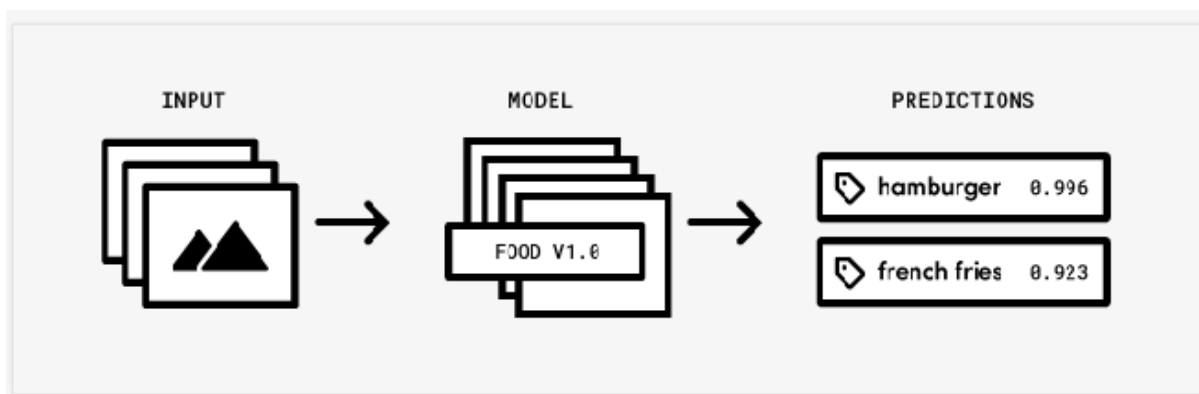


Figure 1: Clarifai API working procedure

2. Expected Inputs and Outputs:

Input	Output
Jpeg Image	Text, audio

Table 2: Expected input and the outputs

IV. Related Work:

Open CV: Open CV is open sourced library which mainly concentrated on the real time computer vision. It supports many image processing operations and it has hundreds of algorithms which can be used to detect the faces, track the moving objects.

It has many programming languages interfaces like C++, C, Python, Java etc., and supports different operating systems as well like windows, Linux, Android.

V. Application Specification

1. System Specifications
 - a. Software Architecture Diagram

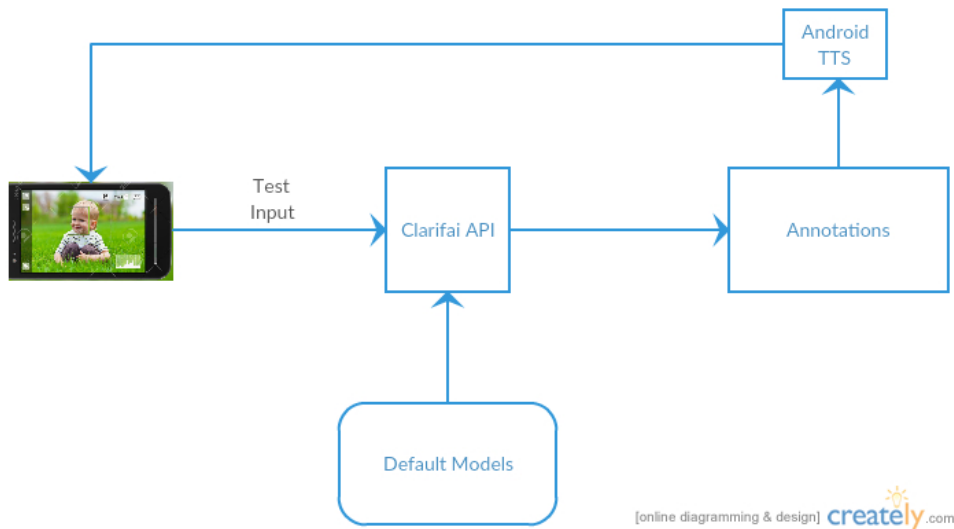


Figure 2: Software Architecture Diagram using Clarifai API.

A test image is fed as input to the clarifai API which is trained with default model and it creates annotation to the test image. The annotation is fed to the Android in built TTS which would convert the text to speech.

b. Activity Diagram:

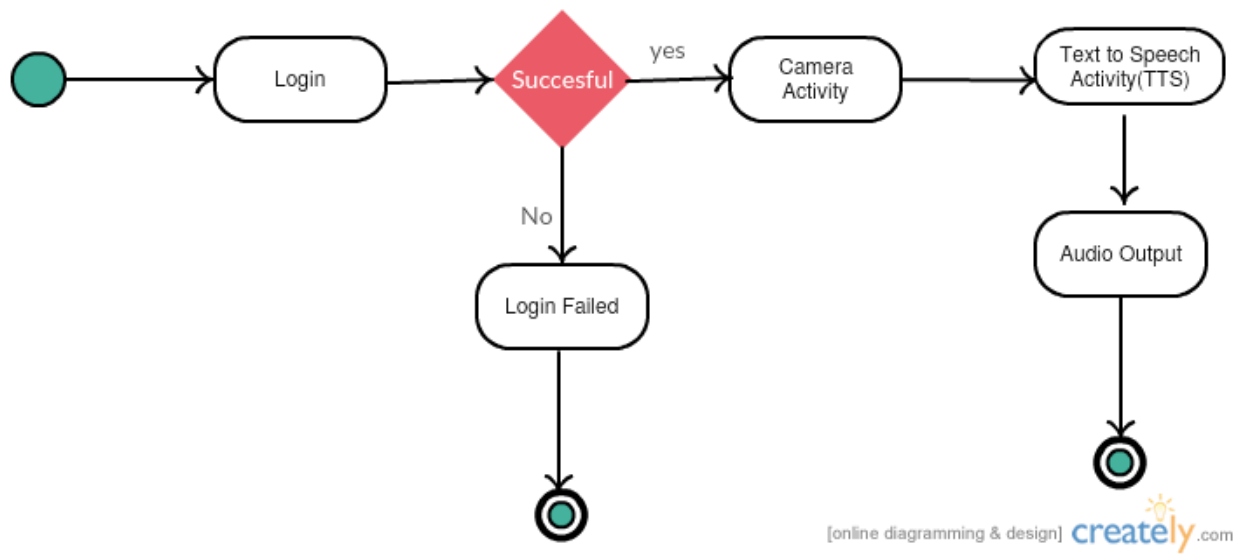


Figure 3: Activity Diagram using Clarifai API.

c. Sequence Diagram

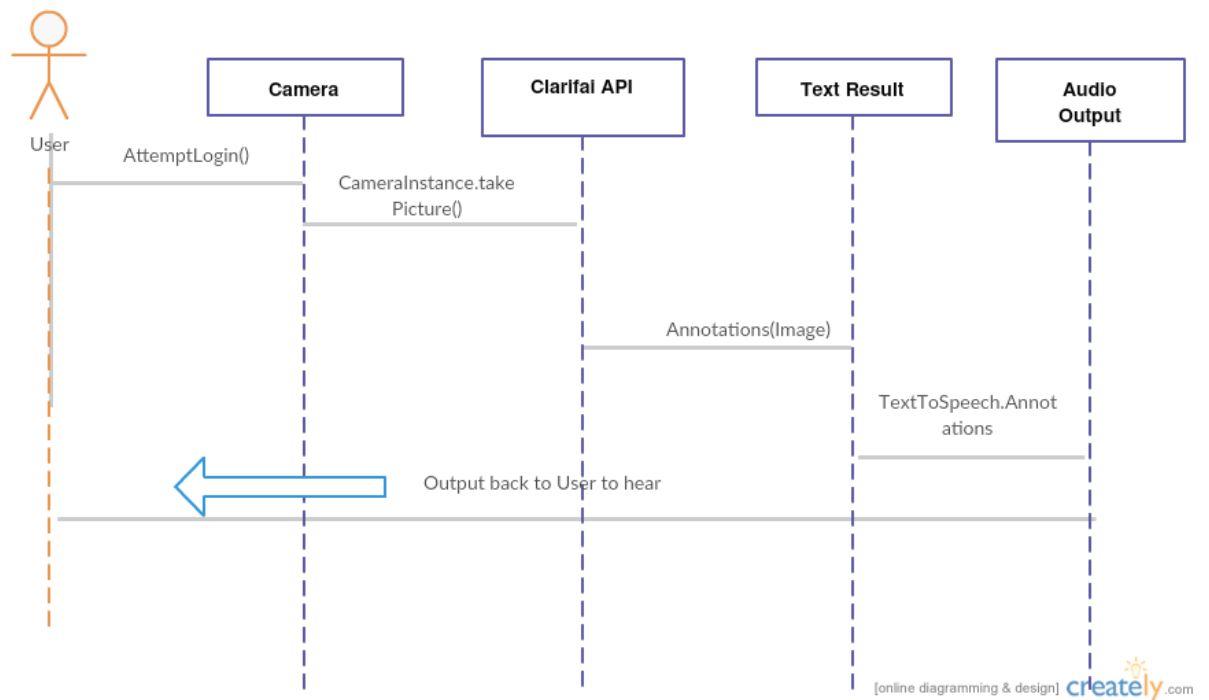


Figure 4 : Sequence Diagram using Clarifai API

VI. Implementation

Client Side: Android app captures video and sends it as input to the server machine.

Server side:

Key frames are generated to the video using Openimaj library and fed as input to the Clarifai or Spark Machine learning API.

Clarifai API:

Takes key frames as input and create annotations to the images.

Spark Machine Learning API:

A model is generated using machine learning techniques using the trained data and the test image is fed as input to classify the image.

VII. Documentation:

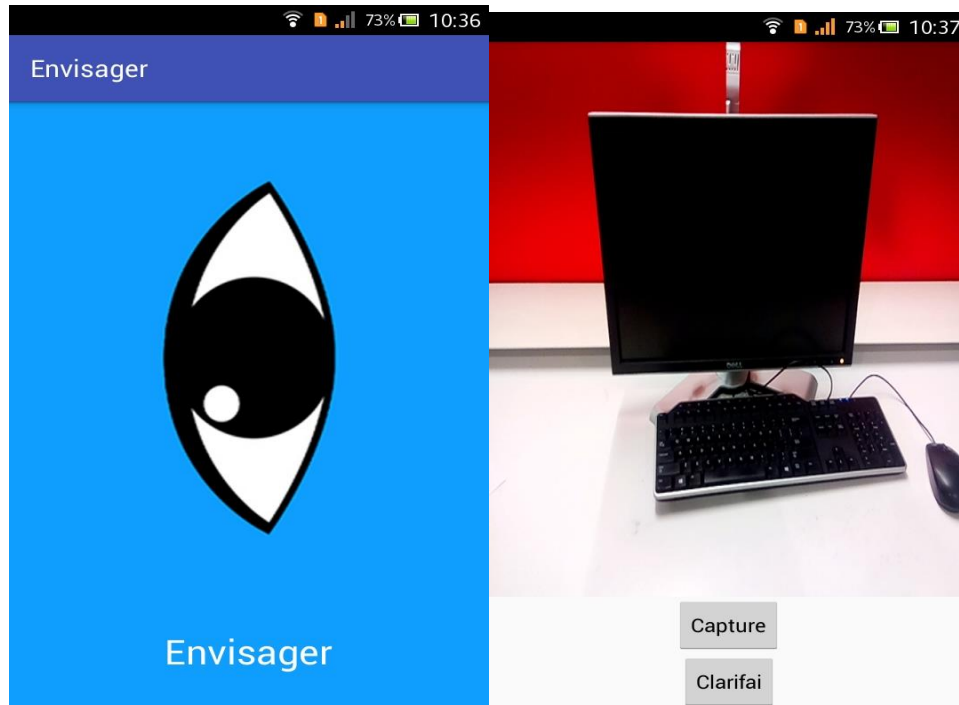


Figure 5: App home page, captured image

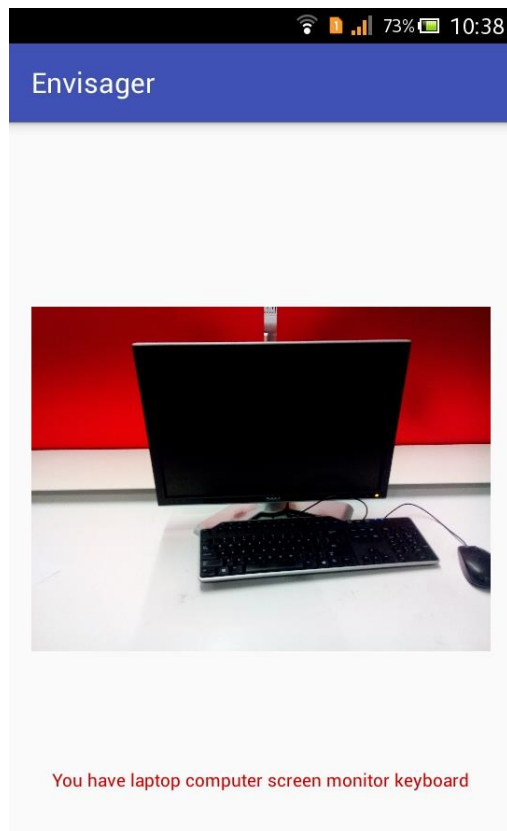


Figure 6: Response annotations

VIII. Project Management:

Implementation status report

Work Completed:

Percentage Contributed	Team Member	Work Done
25%	Venkatesh Gatiganti	Activities, Clarifai API, TTL
25%	Jyothi Kiran Nandanamudi	Activities, Clarifai API, TTL
25%	Madhuri, Gumma	Layout Design, Documentation
25%	Naveena Nallamothe	Layout Design, Documentation

Table 3: Work Completed and Contribution

Project URL:

<https://github.com/nandanamudi/Big-Data-Analytics-and-Application---Envisager>

Work to be completed:

Percentage Contributed	Team Member	Work To Be Done	Estimated Hours
25%	Venkatesh Gatiganti	Client Application for Spark API, Classification,	8 8 8
25%	Jyothi Kiran Nandanamudi	OpenImaj, key frames generation	8 8
25%	Madhuri, Gumma	Text to speech conversion, word2vec,	8 8
25%	Naveena Nallamothe	Make the intellij project executable.	8 8

Table 4: Work to be completed and estimated time.

IX. Reference:

1. <https://developer.android.com/training/basics/firstapp/index.html>
2. <https://www.tutorialspoint.com/android/>
3. <https://classroom.udacity.com/>