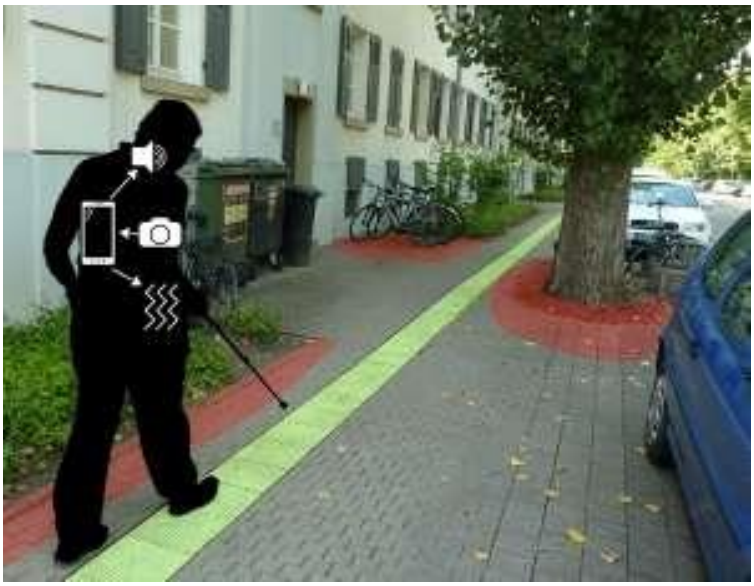
ENVISAGER: SIGHT FOR BLIND AND VISUALLY IMPAIRED

Project Phase - 2



Submitted on **20 March 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)	Android App acting as a client	Built an android app which is acting as a client and sends the request with a test image.

(b)	Image Classification	Used Spark API with Random Forest to classify the test image into predefined classes (labeled classes)
(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 this Phase-2 we developed a spark API to develop a model to classify the test image to a defined class (Supervised learning model).

Input: Image
Output: Audio (Image Explanation)

1. Data Source

ImageNet Dataset: ImageNet dataset has been one of the popular datasets with thousands of images in each node with hierarchy. Even though there are only nouns available for now, there is almost thousand images per node. The database is organized according the hierarchy used by the WordNet.

In in our phase-2 we chose 12 classes to related to our project to train the model and those classes were.

1. Road 2. Vehicles 3. Crossing 4. Cycles 5. Computers 6. Electronics 7. Female, 8 Male 9. Sidewalk 10. Stairs 11. Water 12. Rocks

2. Analytic Tool

Intellij, Android Studio

3. Analytical tasks

Dataset Collection : We used the ImageNet datasource to download the images of a predefined classes related to our project.

Image Scaling : All the images of all the classes were scaled down so that the level of computation will be reduced.

Key Descriptors Extraction: We extracted key descriptors using SIFT (Scale Invariant Feature Transform) which has proven to be a robust local invariant feature descriptor. These features are extracted that are invariant to changes in illumination, image noise, scaling and rotation. As the SIFT features are very resilient to any noise in the image it acts as a good feature extraction descriptor.

Clustering using k-Means: The key descriptors were given as input to the K -means Clustering algorithm(Unsupervised Algorithm). The total observations were clustered into 20 clusters.

Reasons to use this clustering technique:

1. Robust and easy to understand.
2. Relatively efficient.
3. If datasets are distinct, it gives best results.

Histograms generation: The histogram plot is used to show the frequency distribution of continuous image data. This allows to inspect data for underlying outliers, distribution.

This has been a success method for the for bag of words used in image classification by treating the features of the image as words. This bag of words is defined as a sparse vector of the occurrence of the count of words.

Now, the histograms acts as feature vector for every image.

Random Forest Model: Random forest model is a tree which grows with multiple classification trees. After developing a model to classify the input image it stages the image by placing the vectors under each tree in the forest. Each tree proposes the class of the image. Based on the most classified class the class of the image is predicted.

4. Expected Inputs and Outputs:

Input	Output
Jpeg Image	Class the image belongs to, audio saying the class name

Table 2: Expected input and the outputs

5. Algorithms:

1. K-Means : To cluster the descriptor data
2. Random Forest : To classify the image

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, tack 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

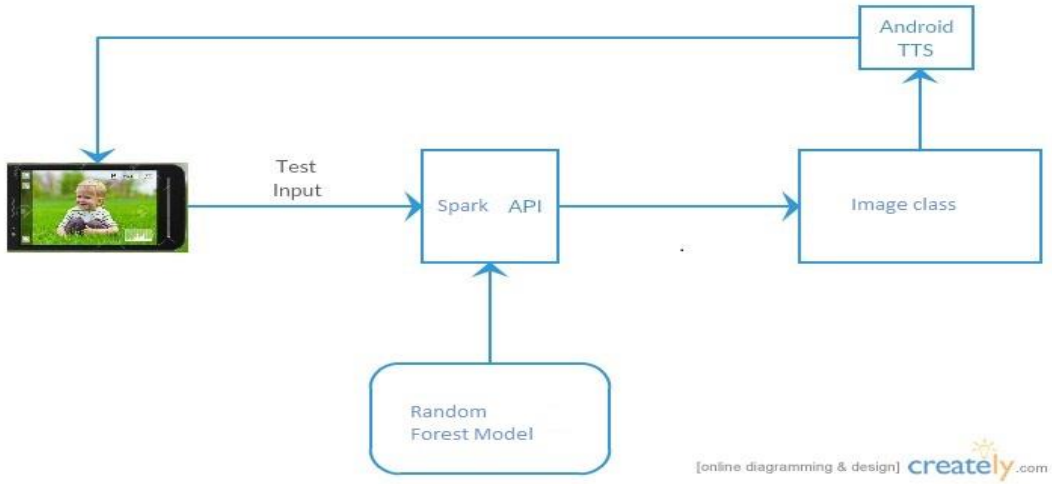
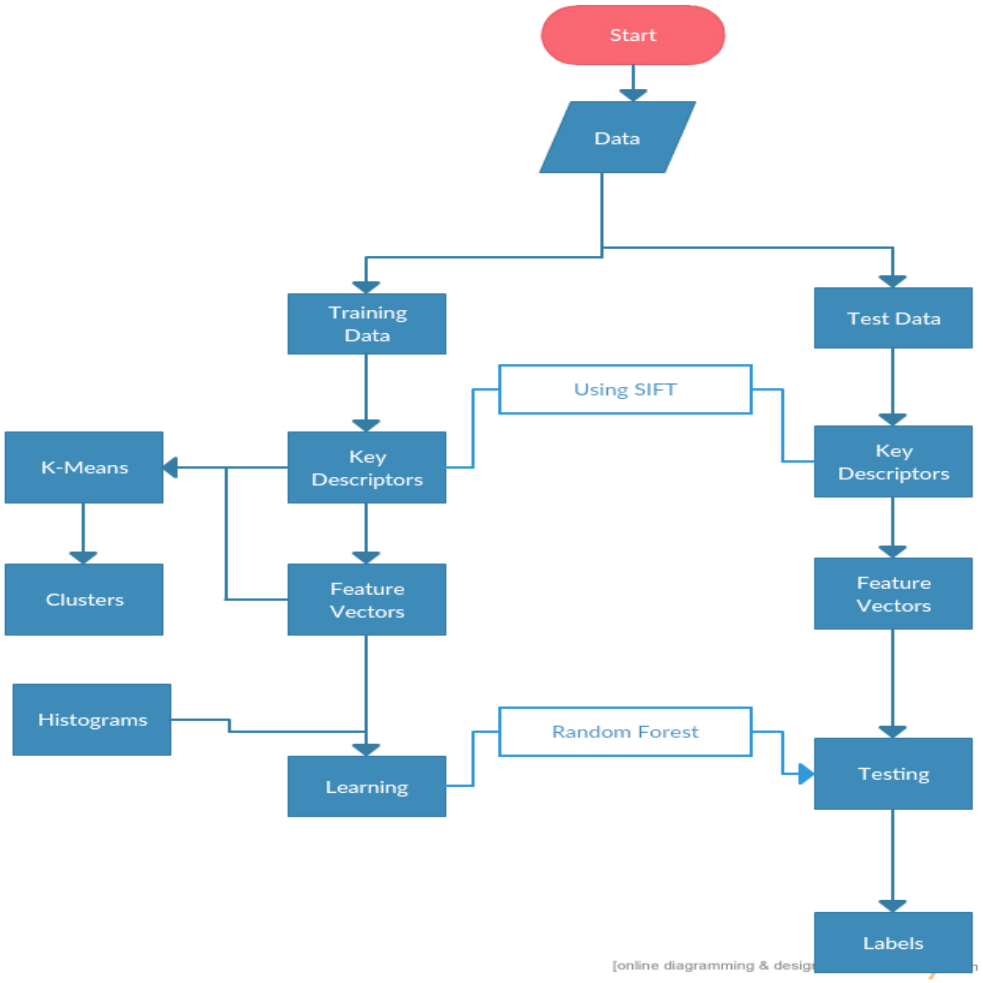


Figure 2: Software Architecture Diagram using Spark API.

b. Flow Diagram:



c. Activity Diagram:

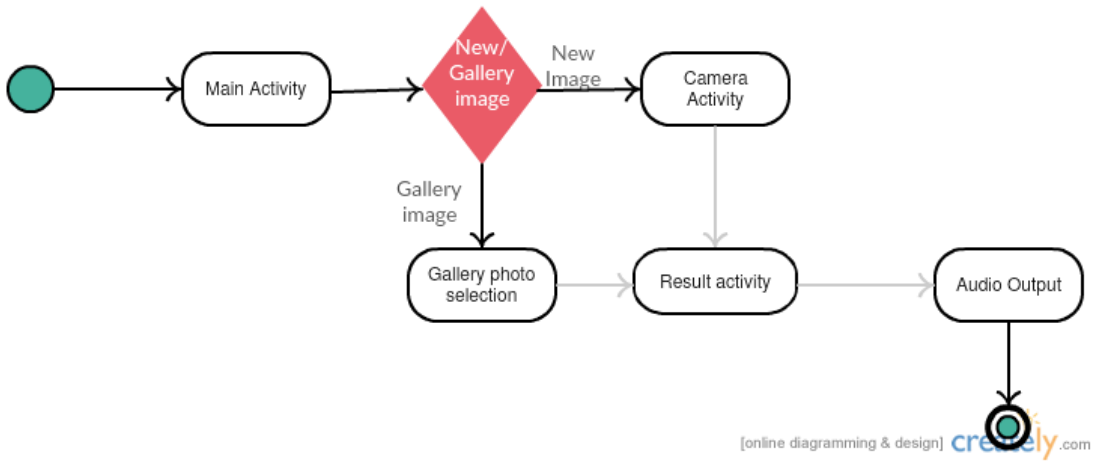


Figure 3: Activity Diagram using Spark API.

d. Sequence Diagram

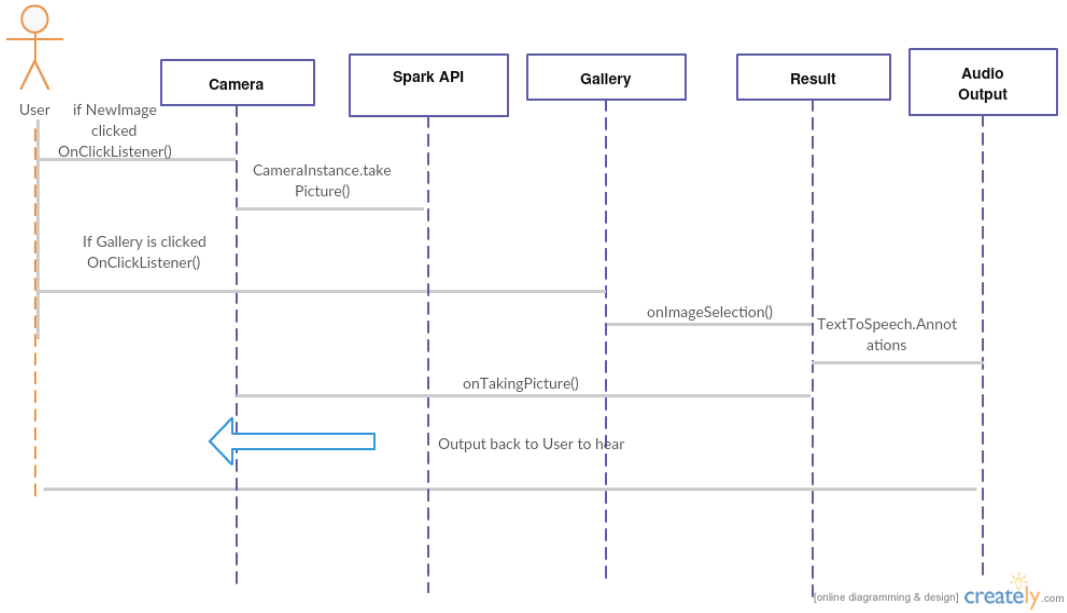


Figure 4 : Sequence Diagram using Spark API

VI. Implementation

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

Server side: The server takes the input as image (test image) and runs with the model already developed with the Random forest. The class of the image is labelled an sent back to the client.

VII. Documentation:

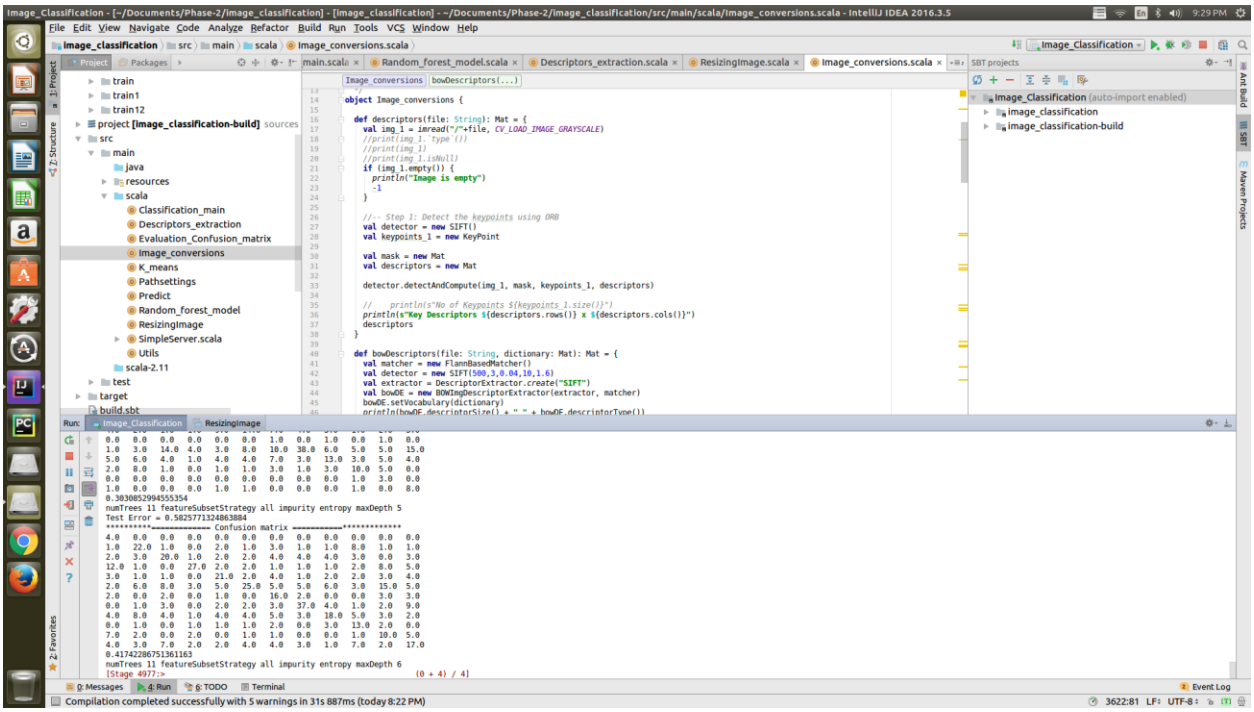


Figure 5: Confusion matrix for each combination in the random forest

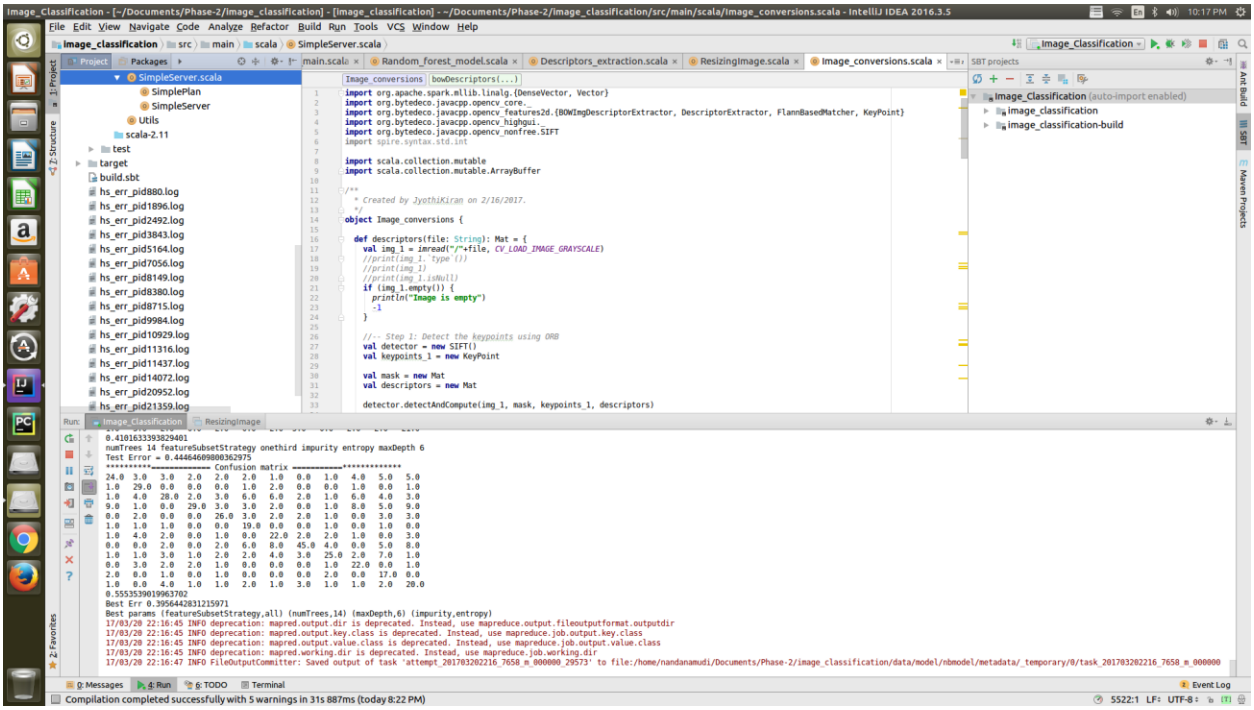


Figure 6: Model Accuracy and Best Parameters for the Random Forest Model

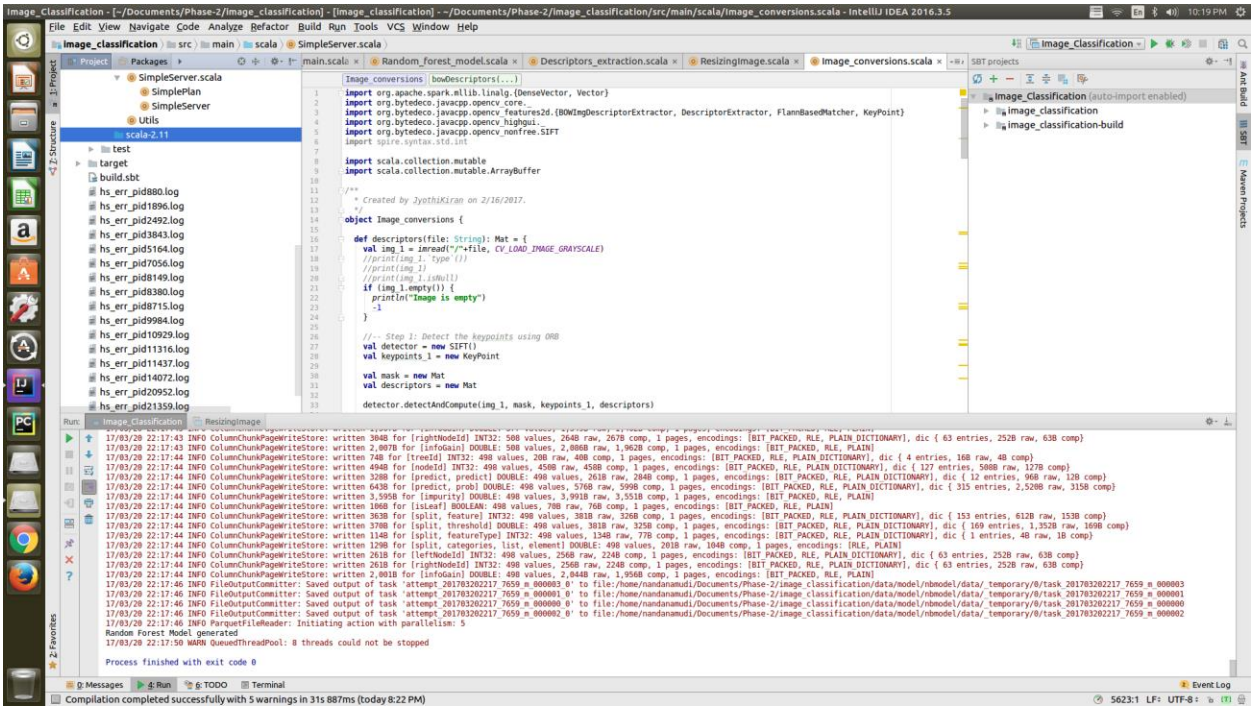


Figure 7: Random Forest Model has been generated

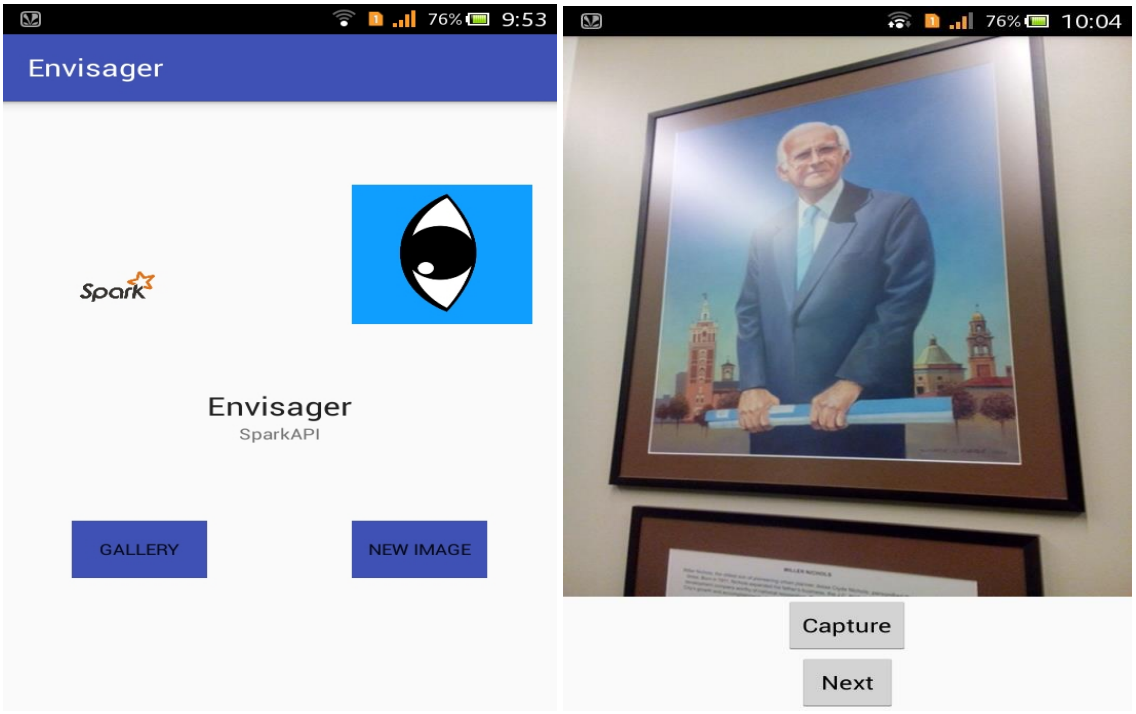


Figure 8 : Home Screen and Image Captured

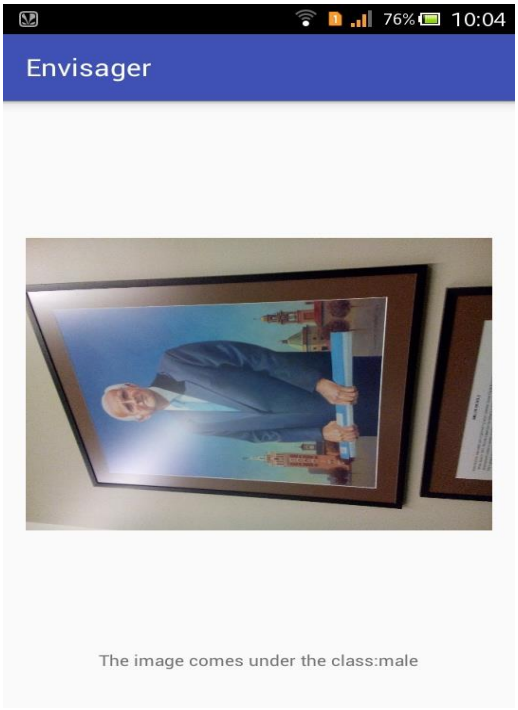


Figure 9: Response: The image is classified under male class.

VIII. Project Management:

Implementation status report

Work Completed:

Percentage Contributed	Team Member	Work Done
25%	Venkatesh Gatiganti	API, TTL, K means clustering
25%	Jyothi Kiran Nandanamudi	Activities, Histograms, RF model
25%	Madhuri, Gumma	Layout Design, Camera Functionality, scaling the images and key descriptors extraction
25%	Naveena Nallamothu	Layout Design, Extracting Data from ImageNet, test image classification

Table 3: Work Completed and Contribution

Project URL:

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

IX. Reference:

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