Landmark Recognition and Retrieval System

Project Id - 18-107

Project Proposal Report

M.M.R.L.K Marasinghe

N.G.K.S Gamage, P.H.R.S Herath, D.A Masachchi

Bachelor of Science (honors) In Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

April 2018

Landmark Recognition and Retrieval System

Project Id - 18-107

Project Proposal Report

(Proposal documentation submitted in partial fulfilment of the requirement for the Degree of Bachelor of Science Special (honors) In Information Technology)

Authors:

M.M.R.L.K Marasinghe (IT14134104)

N.G.K.S Gamage (IT15131416)

P.H.R.S Herath (IT14139604)

D.A Masachchi (IT15033024)

Supervisor:

Mr. Yashas Mallawarachchi

Bachelor of Science (honors) In Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

April 2018

DECLARATION

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
M.M.R.L.K Marasinghe	IT14134104	
N.G.K.S Gamage	IT15131416	
P.H.R.S Herath	IT14139604	
D.A Masachchi	IT15033024	

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

……. …….

Signature of the supervisor: Date

ABSTRACT

Tourist Industry in the World is at a standard stage now. But, still there is no enough push from the technological side for the betterment of the Industry. There are trip planning sites, booking sites and all. But we are in lack of a dedicated online platform for the tourists to identify certain locations. They always in need of a tour guide to do that. But, most of the tourists prefer to travel by their own. They love to face that challenge of being exposed to a different culture by their own. Currently, in the world, there is no systematic approach has been taken for this problem. Because of that we see a gradual reduction of tourists moving out of their comfort zones. This is not good for a country's economy as well. Because tourism plays a main role in the economy of a country. To solve this problem of identifying landmarks, we came up with a mobile based application. This "landmark retrieval and recognition system" helps to ease this task of identifying landmarks by studying the pixels of an image taken by any smart phone device. It doesn't deal with the metadata of the image but it deals with pixel information by using the techniques like image querying. So with that specialty we hope to optimize the landmark recognition process.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1.1 Background

The touristic landmarks are easily recognizable and well-known sites and buildings, such as a monument, church, etc., as shown in Figure 1. They are the pivotal part of people's tours, due to their notable physical, cultural and historical features. The explosion of personal digital photography, together with Internet, has led to the phenomenal growth of landmark photo sharing in many websites like Picasa Web Album (picasa.google.com). [1]

Figure 1: Examples of landmarks in the world

Picture recuperation is a focal issue in Computer vision: given a request picture, would you have the capacity to find similar pictures in a large database? This is especially basic for question pictures containing purposes of intrigue, which speaks to a colossal portion of what people get a similar out of the opportunity to photograph.

With the huge measure of milestone pictures in the Internet, the time has wanted Computer vision to consider historic points all around, in particular to assemble a point of interest acknowledgment system, on the size of the whole globe. This system isn't just to outwardly perceive the nearness of specific landmark points in a picture, yet in addition adds to an overall milestone database that sorts out and records

historic points, regarding land areas, popularities, social esteems and social capacities, and so forth. Such an earth-scale point of interest acknowledgment motor is enormously helpful for some vision and mixed media applications.

1.2 Research gap

We are in a need of well-designed solution that would allow us to use a single standardized System for Landmark recognition and Retrieval in World. In the current available system can identify and give the location. They used Google Map by using mobile devices for this purpose. Also they divide segments in input image compare their database and show them what that landmark is.

In current system like Google Goggles – Search by taking a picture: point your mobile phone camera at a painting, a famous landmark, a barcode or QR code, a product, or a popular image. If Goggles finds it in its database, it will provide you with useful information [2] (name, country)

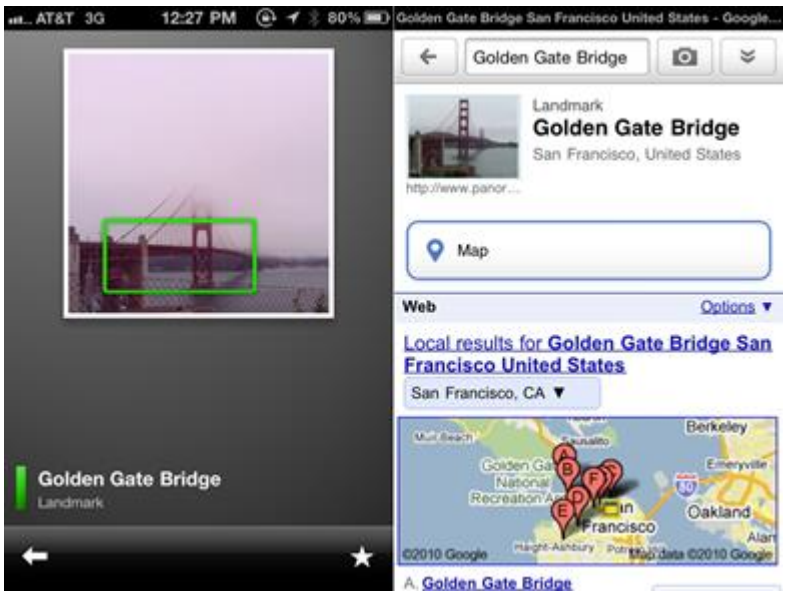

Figure 2: Finding Goggles

Present system does not have recognize the landmark images by using pixels. They only compare that input image is same as there data base image and retrieve it. Also they are not Consider several human features. From the survey that has to be filled by the user and predict the places according to that particular area. We mainly focused on those issues and came up with better solution for that system with high reliability, better accuracy.

1.3 Research Problem

As mentioned earlier, there are major problems with the current system. When we see some images of place that we are interest of visiting or go for some more images on that same landmark but the image retrieval is fundamental problem in computer vision make a barrier to retrieve the image with correct landmark from large databases.

Also if we have any collection of images where we visited while ago but no extract memory right now like what is the name of this temple I visited? Who created this monument I saw in there? And also we are curious of the more details on that monument or the temple. Build models that recognize the correct landmark (if any) in a dataset is challenging task. Then Consider several features from the survey that has to be filled by the user and predict the places according to that particular area. [3]

1.4 Literature Survey

Today the people are eager to travel all round the world and the information technology has involved facilitating them by providing the information about the places they are willing to explore.
There are many mobile applications have been implemented to reduce the difficulties they face when travelling and organizing the memories. Most of them have considered the metadata for this purpose while we are going to use the content based data for the recognizing and retrieval as it is more accurate. The accommodations and the real time update of the place. And for the person who is going to upload an image to the application, they can get a feedback of others too. Most importantly this data exchange is encrypted

Project 01: A Visual Landmark Recognition System for Topological Navigation of Mobile Robots [4]

This research describes a vision-based landmark recognition system for use with mobile robot navigation tasks. A search algorithm based on genetic techniques for pattern recognition in digital images is presented. The developed system allows the topological localization of a mobile robot using natural and artificial landmarks. Text strings inside landmarks can be read and interpreted, if present. The resulting system has been tested onboard a B21 mobile robot and proved useful. The presented experimental results show the effectiveness of the proposed algorithm. [4]

Project 02: Fast Online Learning Algorithm for Landmark Recognition based on Bow Framework [5]

It proposes a quick web based learning system for Landmark acknowledgment in light of single shrouded layer feed forward neural systems (SLFNs). Ordinary Landmark acknowledgment structures by and large accept that all pictures are accessible within reach to prepare the classifier. Nonetheless, in certifiable applications, individuals may experience the issue that the classifier based on the current historic point dataset should be tuned when new milestone pictures are gathered. To address this issue, a quick online consecutive learning structure in view of the current extraordinary learning machine (ELM) which can refresh the classifier by taking in the new pictures one-by-one or lump by-piece is produced for the point of interest acknowledgment. The current spatial pyramid kernel bag-of-words (bow) technique is utilized for the element.

Project 03: City-Scale Landmark Identification on Mobile Devices [6]

With recent advances in mobile computing, the demand for visual localization or landmark identification on mobile devices is gaining interest. We advance the state of the art in this area by fusing two popular representations of street level image data—facade-aligned and viewpoint-aligned and show that they contain complementary information that can be exploited to significantly improve the recall rates on the city scale. We also improve feature detection in low contrast parts of the street-level data, and discuss how to incorporate priors on a user's position (e.g. given by noisy GPS readings or network cells), which previous approaches often ignore. [6]

Project 4: RMI for mobile behavior

RMI can move behavior (class implementations) from client to server and server to client. For example, you can define an interface for examining employee expense reports to see whether they conform to current company policy. When an expense report is created, an object that implements that interface can be fetched by the client from the server. When the policies change, the server will start returning a different implementation of that interface that uses the new policies. The constraints will therefore be checked on the client side-providing faster feedback to the user and less load on the server-without installing any new software on user's system. This gives you maximal flexibility, since changing policies requires you to write only one new Java class and install it once on the server host.

	Project 1	Project 2.	Project 3	Project 4	Our Project
Landmark Recognition component					
Landmark Retrieval component					
Predict user reservations by profiling					
Distributed Encrypted data among the System					

2. OBJECTIVES

2.1 Main Objective

Provides a systematic solution to predict the landmark details of the images and a way to find (query) images containing the same landmarks to help people better understand and organize their photo collections. [3]
Also the system predict the accommodations and dining suitable for the customer by his preferences and the real time update of the place they are willing to explore.

2.2 Specific Objectives

- Investigate existing problems in landmark recognition and landmark retrieval in computer vision Retrieval.
- Determine the possible solution or best methods to address those challenges.
- Build a model (algorithm) to recognize the correct landmark of an image
- Build a model (algorithm) to find similar images in a large database
- Concurrently access and view particular images recently updated.
- Generate the rating Landmarks & Countries.
- Improve the user experience by user profiling
- Provide user-friendly interfaces and easy access to the system.

3. RESEARCH METHODOLOGY

3.1 System Overview

The proposed system consists of a mobile application mostly focused on tourists and it will be even helpful for the tour guides. Mobile application interacts with travelers who likes to find the details of the particular image details that contain landmarks.

This application is connected via cloud server to connect with inbuilt Restful server in our application. In addition to that it implements an image processing system to capture the landmarks to identify and gather the details of the landmark by using OpenCV library and python. Also the system is capable of customizing the user experience by using the profiling technique we implement with Twitter.

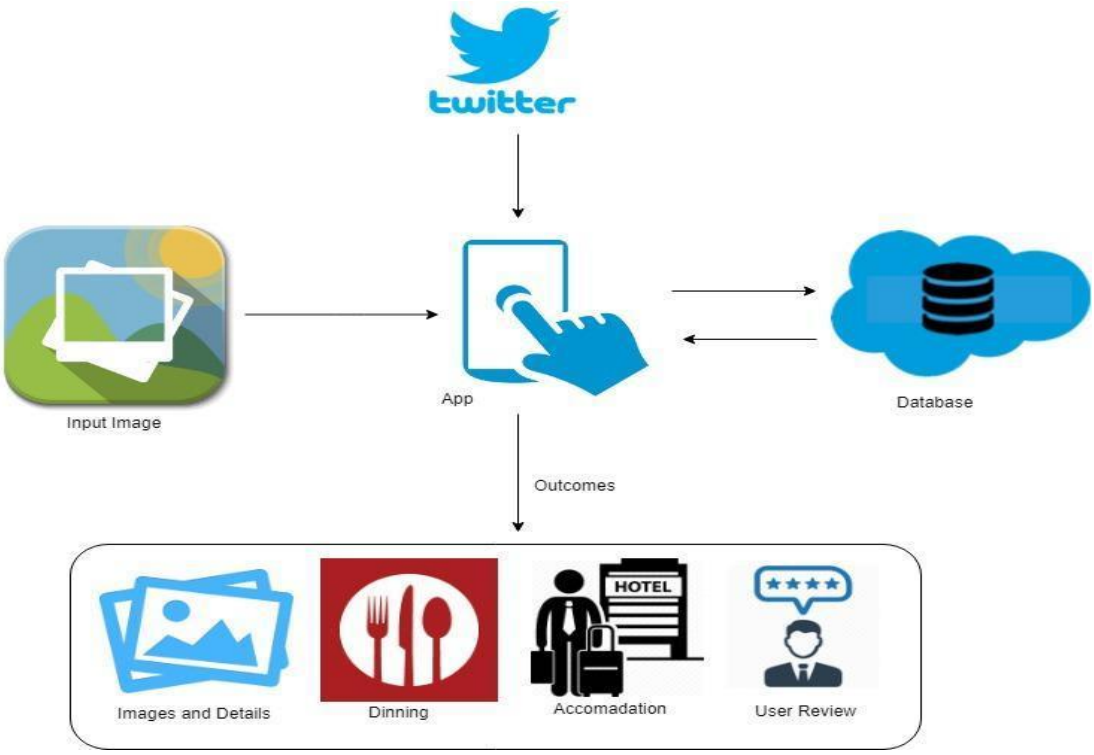

Figure 03: System Overview

3.2 Functionality of the project Main Components

3.2.1. Landmark Recognition component

This component will be there take test images and identify which landmarks (if any) are depicted in them. This Test images may depict no landmark, one landmark, or more than one landmark. [3]

This component is responsible for:

- Analyze the upload/captured image by using image processing techniques to identify whether it is contain a landmark or not.
- Storing and retrieving landmark information such as name, location, etc. from the captured image by using image processing and machine learning techniques.
- Upload the captured image to cloud server as a background process to identify the landmark using image processing and machine learning techniques.
- Build a machine learning algorithm that recognize the correct landmark.

Main objective of this component will be find a way to recognize correct landmarks in a large dataset of challenging test images.

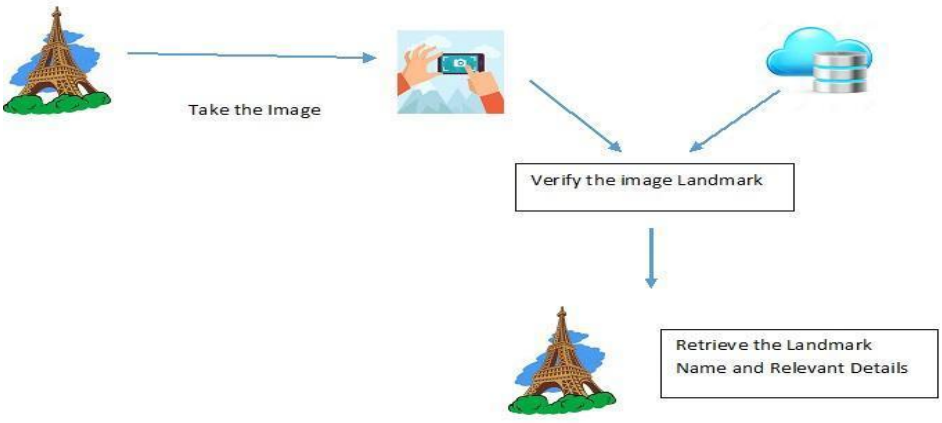

Figure 4: Overview of landmark Image Recognition

3.2.2. Landmark Retrieval component

The image retrieval is a fundamental problem in computer vision. Considering the content based image recognition to find the same images captured in that landmark.

Once an image is uploaded, the system will use that image as an input to use under image processing technology. As a result, the system will check that result through the image database and the system itself will provide similar images or duplicate regarding that image which was uploaded as an input. [7]

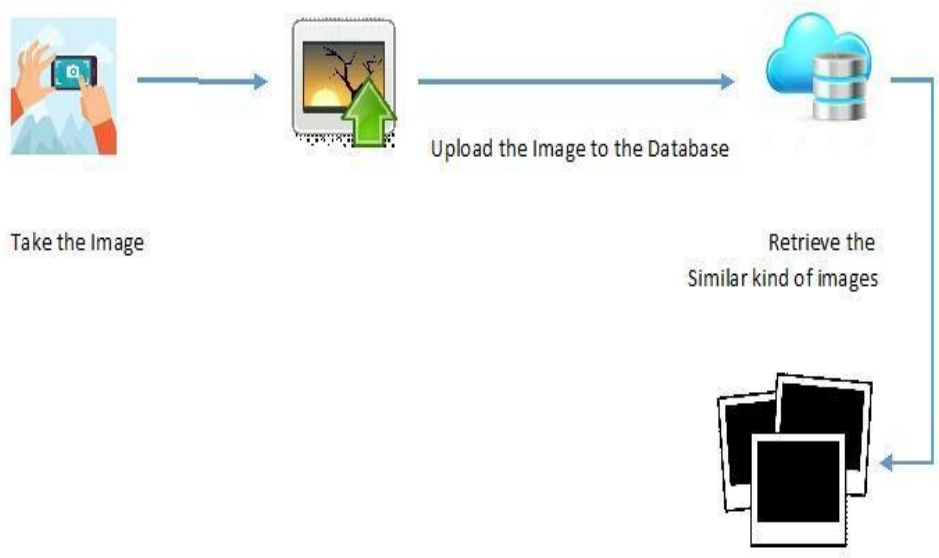

Figure 5: Overview of Image Retrieval

3.2.3. Predict user reservations by profiling

User profiling is used to personalize the online experience of the user while using the app. Twitter provides an API to access the publicly available information. The words, Hashtags, Retweets etc. can be accessed from the user profile and they can be used for the profiling.

Twitter offers the data through two channels, the Stream API and the REST API. By using the python package called Twython we can connect the both APIs to collect data. Also we use GIS technology here to get the location of the captured image of the landmark and it will be further used to provide the user with Restaurants and dining places around that particular area. But, the results will be displayed according to the user preferences we obtained through Profiling. For example, user can go directly to the restaurant which offers his/her favorite food rather than going to every restaurant and check for their favorites.

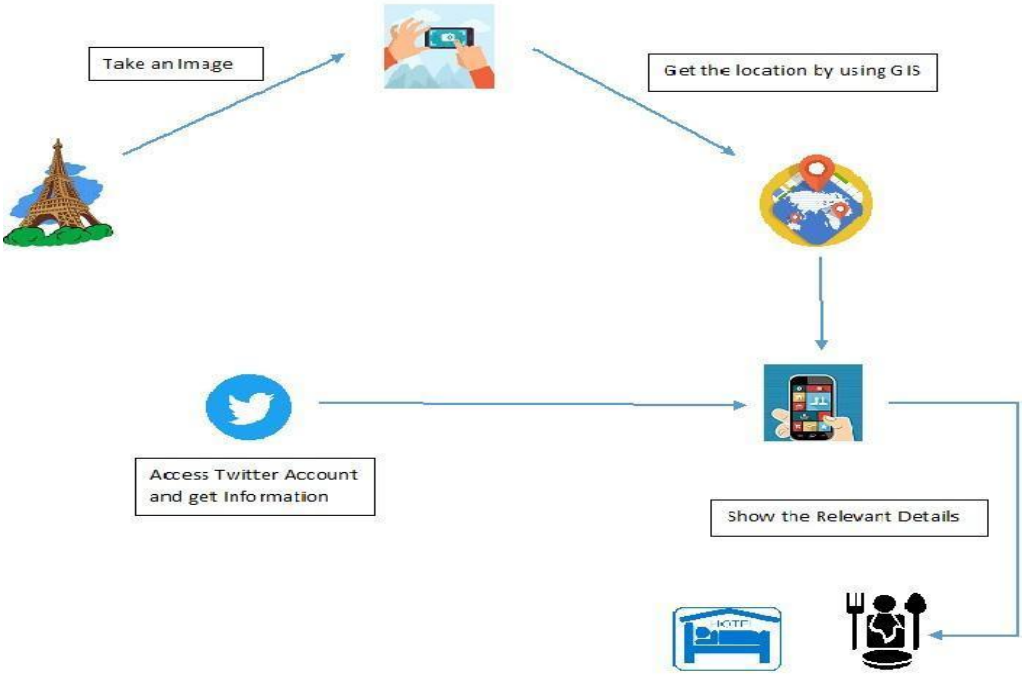

Figure 6: Overview of profiling

3.2.4. Distributed Encrypted data among the System

Main Objective of this component is to deliver and process encrypted data that receives from the domains. When someone take image and upload concurrently then can show that image who use that mobile app. HTTPS will be used so that the web application will be more secure. Then user can see number of views that particular image. RMI Systems are used for that real time updates.

Therefore, the security of distributed data would be safe and also the process time would be reduced.

In this component we can put the rating and identify what are places most people visited.

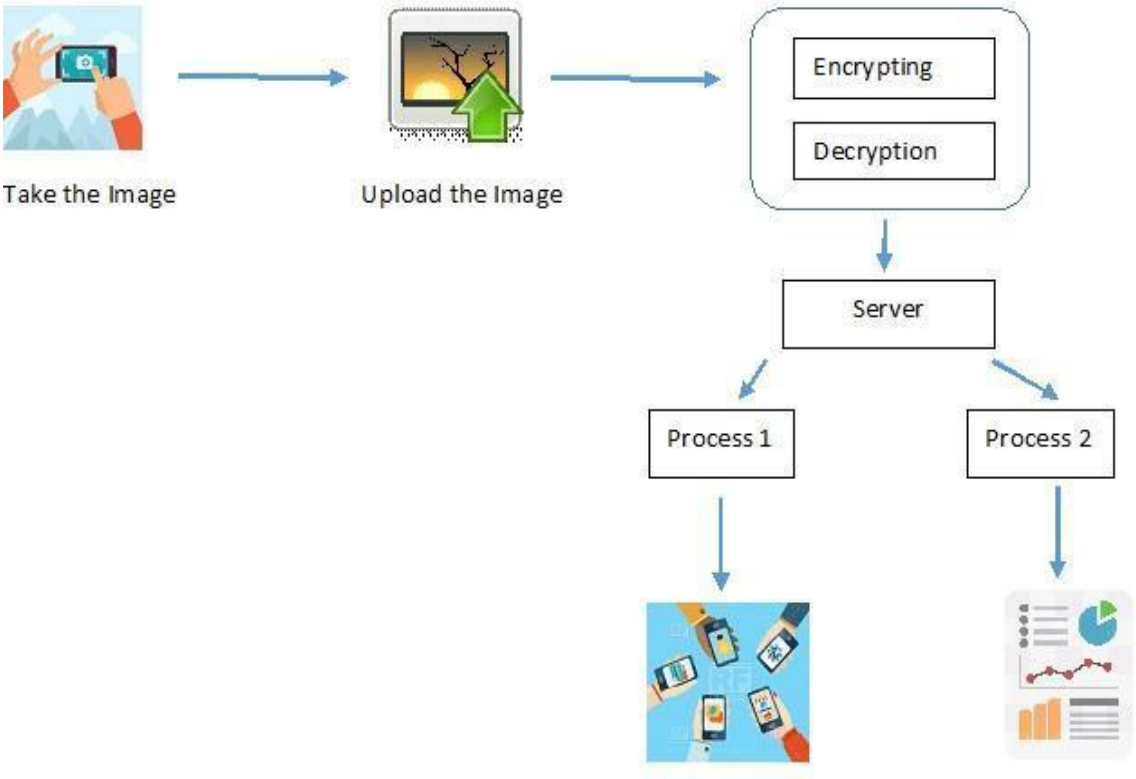

Figure 7: Overview of Distributed Encrypted data

3.3 Technical Requirements

Hardware Requirements

1. Android support mobile device (Marshmallow or above mobile operating system)

2. At least 600MB internal space Mobile device.

Other Requirements

1. Cloud Base Database - Firebase

2. Algorithms – Viola Jones Object Detection Framework

YOLO

K-means Algorithm

Support Vector Machines Algorithm

4 .Mobile application development technology

5. RMI Distributed Systems for Real Time Updates

6. OpenCV library and python.

7. MATLAB

8. Stream API

9. REST API

10. Twython (To connect to the both APIs)

3.4 Project Outcomes

Outcome of this project is to build an application for tourists which will help them in identifying the landmarks by using the captured image of the particular landmark. This research will introduce new platform for landmark recognition with image querying by identifying the pixel information of the image. It will replace the traditional metadata analysis of an image. Users don't need any other special device for this. They just need their smartphone to capture the image and internet access to connect to the service. Using the application tourists/backpackers can identify a landmark accurately and they can get information about that particular landmark. And also users can search for the similar images of the same landmark or can get information about other places of the same domain. For example, if you clicked at an archeological site the app can give you information about other archaeological sites in that area. Plus, with a linked twitter account the user profiling can be done. Likewise, it captures the user preferences like interested tourist places, food preferences etc. from the twitter user profile or timeline. So when the user travels to a certain place, the app can suggest restaurants around that area which serves his/her favorite food types and preferred places for accommodation. Also the users can write reviews about the places they visit, so that other tourists will also can get the benefit from those reviews.

3.5 Gantt chart

Figure 08: Gantt chart

4. Marketing the application

Through our app the user will be able to get information about a landmark with real time updates about the place including user reviews. Also through our app the user is able to get a customized experience with his/her preferences which we will be obtained through User profiling technique. For an example, a user can find a restaurant with his/her favorite foods available. Plus, with the GIS technology the user will be getting that data specified to the current location or to the specified location of the landmark. For the convenient of the user we have implemented the reservation feature also in to the app.

5. DESCRIPTION OF PERSONAL AND FACILITIES

Member	Component	Task
Marasinghe M.M.R.L.K. IT14134104	landmark retrieval	• Recognize the query image • Build a model to Retrieve the images that match for given query image
Gamage N.G.K.S. IT15131416	landmark recognition	• Build a model to recognize the landmark o Collecting Labeled Data o Analyzing Data o Splitting the Data o Training the Model • Retrieve correct landmark details
Herath P.H.R.S. IT14139604	Predict user reservations by profiling	• Access data from Twitter by using the channels Stream API and the REST API • Artificial data model validation
Masachchi D.A. IT15033024	Distributed Encrypted data among the System	• Concurrent Accessing • Display Number of views particular image • Generate the Rating • Identify most visited Landmarks & Countries.

6. BUDGET AND BUDGET JUSTIFICATION

No of days spent for the project	=	200
Average working hours per day	=	3 hours
No of people engaged for development	=	4 people
Labor hours per industry	=	Rs.30 per hour
Total working hours	=	3 * 200* 4
	=	2400 hours
Total working cost	=	2400 * 30
	=	Rs.72, 000.00
Travelling Expenses	=	Rs.15, 000.00
Printing and stationary charges	=	Rs. 6,000.00
Cloud based services	=	Rs. 8,000.00
Total cost	=	Rs. 101,000.00

REFERENCES

[1] Y. T. Zheng *et al.*, "Tour the World: Building a web-scale landmark recognition engine," *2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009*, pp. 1085–1092, 2009.

[2] D. Petrou *et al.*, "Search by sight: Google™goggles," *2010 IEEE Hot Chips 22 Symp. HCS 2010*, 2016.

[3] Google, "Google Landmark Recognition Challenge | Kaggle." [Online]. Available: https://www.kaggle.com/c/landmark-recognition-challenge. [Accessed: 20-Mar-2018].

[4] M. Mata, J. M. Armingol, A. De la Escalera, and M. A. Salichs, "A visual landmark recognition system for topological navigation of mobile robots," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 2, pp. 1124–1129, 2001.

[5] J. Cao, T. Chen, and J. Fan, "Fast online learning algorithm for landmark recognition based on BoW framework," *Proc. 2014 9th IEEE Conf. Ind. Electron. Appl. ICIEA 2014*, pp. 1163–1168, 2014.

[6] D. M. Chen *et al.*, "City-scale landmark identification on mobile devices," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 737–744, 2011.

[7] Google, "Google Landmark Retrieval Challenge | Kaggle." [Online]. Available: https://www.kaggle.com/c/landmark-retrieval-challenge. [Accessed: 20-Mar-2018].

[8] N. Dalal and W. Triggs, “Histograms of Oriented Gradients for Human Detection,” *2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR05*, vol. 1, no. 3, pp. 886–893, 2004.