Monument Recognition

Kumar Satya Gopal Panja, Sivarama Krishna Linga, Vikesh Padarthi

University Of Missouri Kansas City, MO 64110, USA

Abstract

We have developed application to do image analysis by using existing algorithms and by modifying them so that these algorithms would give us the efficient results. As far as this project is considered it accepts the image and does analysis on that image and gives us the data about the image and related things. We evaluate data which we obtain after analyzing those images. We have collected almost 10k set of images and all these images are related to historical monuments. We have trained the model using these images and generated key descriptors while implementing the analysis using the machine learning algorithms. We have implemented image classification and clustering. For this we used algorithms like Naive Bayes, Decision tree etc., we have used all these algorithms to know which algorithm gives more accurate results for our project. Not only by using Machine learning algorithms we have implemented but also using Deep learning concepts we have done this. Although using Tensor flow gives us some accurate results but in our case, we came to conclusion that using the machine Learning will give us some good results and it would be easy to handle. user can take a picture from his mobile of a place or a building and he can upload it to the application and he can learn about the details of not only the image i.e. its history like when it was constructed, who constructed it, during which time it was developed etc. We are also giving the data like places which are nearer to him which he can see with in some area and timings when they are open etc. In this way it would be helpful for the user to make his plans. This is the idea of the project and we have implemented most of this.

1.INTRODUCTION:

The main idea of developing this project is to help the people to learn about the places which they love. Suppose when we see some random pictures we never know from where they are. These images might be buildings or historical monuments or some famous streets. Some people will love to visit those places but as they don't have any idea they get stopped thinking about that.

So to overcome this and help the people to learn about the pictures we have developed this project. To completely achieve this functionality it really need huge amount of data set because there are more than like thousands of historical places around the world and if we want to consider famous statues there are like hundreds of them. So the training data is huge. Whenever people upload the image they will get the description about the image and not only that but also recommendations.

Not only we have used Machine Learning algorithms but also used Wikipedia API to get the details about the image. For example, when the people upload the image of Eiffel Tower or Taj Mahal, they will get the name and also fewer places they can visit in that place. For Taj Mahal, they will get recommendations like Indian Gate, Red fort etc.

While browsing the internet and Facebook sometimes we used to see different images like buildings, statues etc. But when we try to figure out what they are and where they are located we couldn't find any. Although we use Google Search, it is not accurate as we expected. So We thought that this application i.e. developing an application would really helpful where it would take just image as an input and give us the result.

We want to extend this by uploading the videos too. We know that in recent times it is getting easier to take 360 degrees video. people love those kinds of

videos. So when a person uploads a video he or she should be able to learn about the video.

As mentioned previously, the main goal of the project is not only recognizing the image but also suggesting the places to the user. For this application, we need not to have complex algorithms but the main thing is the training data should be huge and it should be able to match the images accurately.

2.RELATED WORK:

From our Research and analysis, we couldn't find any websites or applications which would do image analysis and help us to find the details. We know that there is Google image search which would be helpful to learn about a village. But google image search is not that much accurate for some of the images. So, we thought this would be great idea to develop the image searching algorithm which would help to learn about the image.

Although it would help the people at the same time it requires lot of images to be trained. This is the only disadvantage of this model. Except that it has everything which would help the people to learn about the images.

3. Proposed Work:

3.1 Proposed Models:

Here we design monument recognition using three techniques.

1. API based visual question answering system
2. Machine learning based visual question answering system.
3. Deep learning based visual question answering system

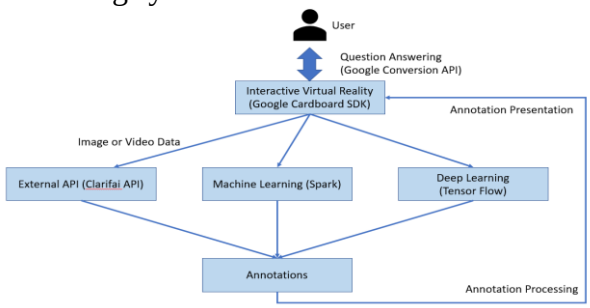

Fig 1

3.1.1.API based question answering system:

Here we used Clarifai API for visual question answering system. The Clarifai API provides image and video recognition as service. It takes images as input and gives predictions as output. We used the our own monument dataset for detecting what is present in the image.

3.1.1. Machine learning based question answering system:

Here we designed question answering system based on several machine learning algorithms and spark techniques. Mainly we used scala for this project. By using REST services, we connected

3.1.2 Deep learning based question answering system:

Here we designed web based visual question answering system using tensor flow. By using different deep learning algorithms and python language we completed this system.

3.2 Algorithms:

There are several algorithms that can be used for analyzing the data sets. However, based on the application appropriate algorithms need to be chosen. For example, Naïve Bayes performs better when the data set is a text data. On the other hand, image data gets better results when k-means is used. Hence, we used K-means for classifying the image. Initially the number of clusters and iterations are predefined and the given dataset is clustered into different clusters based on the common attributes. Then the model is built based on train data. And we used different deep learning algorithms. We used SoftMax, Inception concepts here.

4. Implementation:

4.1 System Design & Implementation

4.1.1 Software Architecture:

Below image clearly specifies how we have implemented this project. Each and every step was briefly explained in this image.

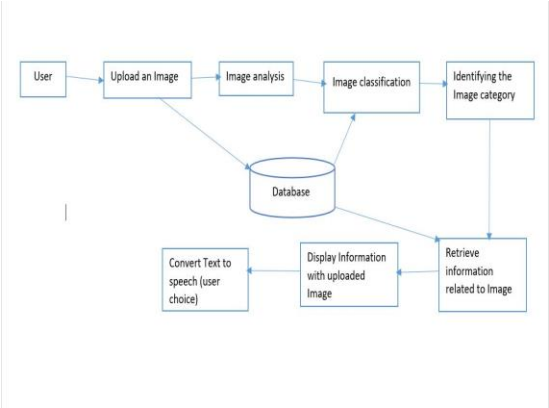

Fig 2

If we could see the above image Fig 2, it is clear that the flow of the project goes exactly as mentioned above.

4.1.2 UML Implementation

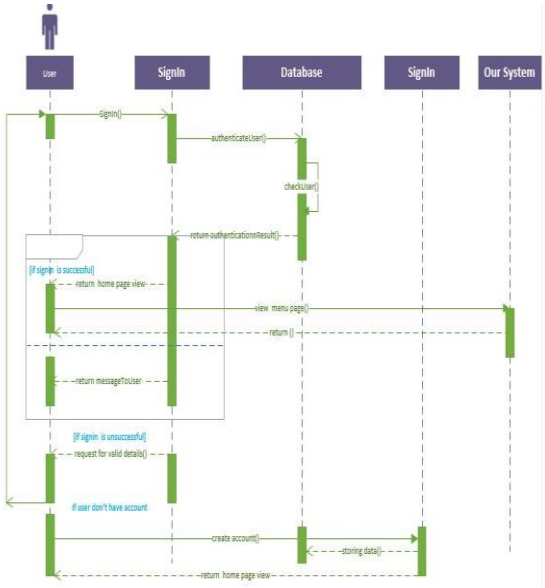

Fig 3

The Fig 3 explains us the UML structure of the project and other details.

4.1.3. Implementation:

As part of Implementation, we have done the project using clarify API and from the below images you will find the output of those implementations.

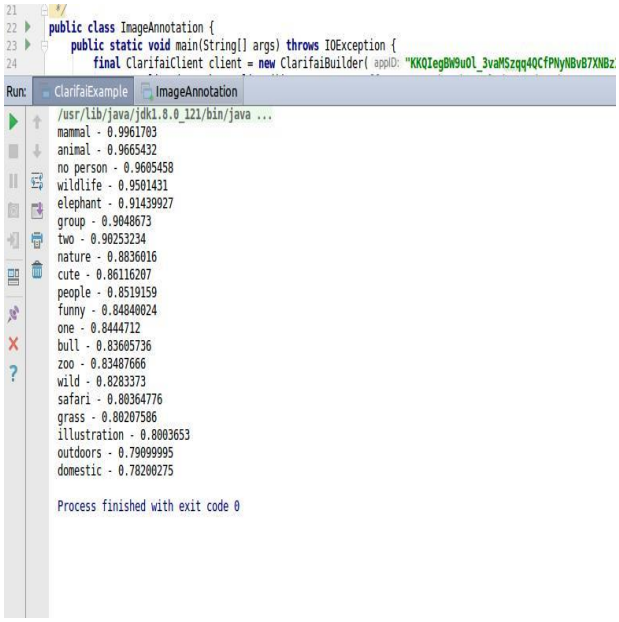

Fig 4.a

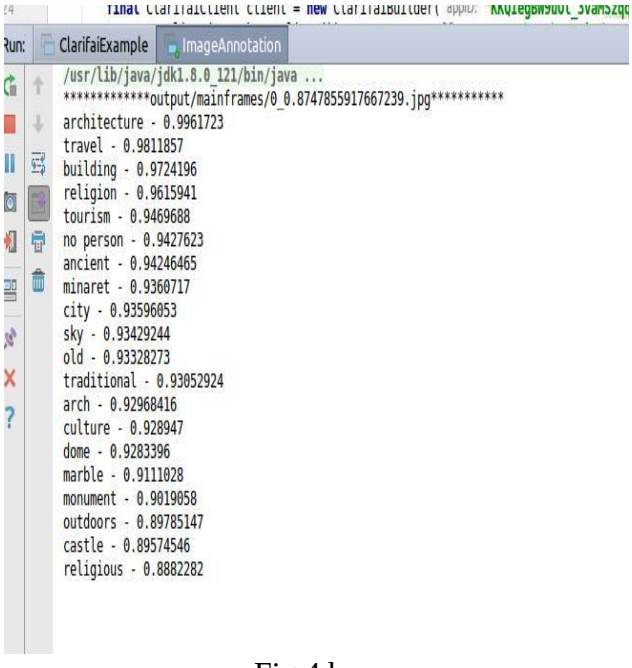

Fig 4.b

And we designed viual question answering sysem by using machine learing techinques. It looks like as follows:

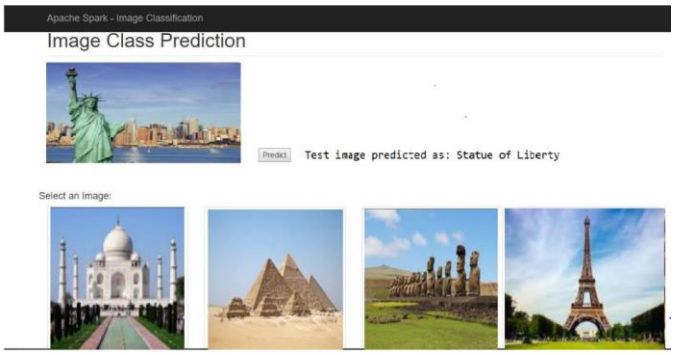

Fig 5

And we designed visual based question asnwering system. It looks like same as above one, but we used different tensor flow techniques here.

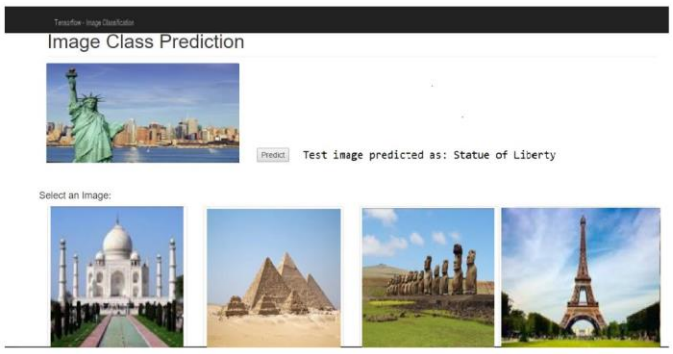

Fig 6

5.Evaluation & Results:

5.1 Dataset:

Here we used our own data set. It contains both train data set and test data set. It contains 9 classes. And each class contains around 15-20 images. And we divided the data set into train data set and test data set.

5.2 System specification:

Operating System: Ubuntu
Minimum RAM Size: 8GB (we can use less than 8 GB but it takes so much time)
Languages Used: Java, Python, Scala
HTML, CSS, JS, Node.js (for front end)
Software Used: Intellij, Pycharm

5.3 Measurements:

Here we take accuracy as main measurement. We got 98% accuracy by using tensor flow.

5.4 Results:

Here we designed same application three times by using different techniques. These three gives better results, but we got superior results using tensor flow. And based on train data and test data size, accuracy varies. We found confusion matrixes for deep learning and machine learning algorithms.

```
5.0  0.0  1.0  0.0  1.0  0.0  1.0  0.0  0.0
3.0  4.0  1.0  1.0  2.0  0.0  1.0  2.0  0.0
0.0  0.0  7.0  3.0  0.0  1.0  1.0  1.0  0.0
0.0  0.0  0.0  4.0  0.0  0.0  1.0  0.0  0.0
1.0  0.0  1.0  0.0  3.0  0.0  0.0  0.0  1.0
0.0  0.0  2.0  1.0  1.0  3.0  0.0  0.0  0.0
0.0  0.0  0.0  0.0  0.0  0.0  3.0  0.0  0.0
0.0  0.0  1.0  0.0  1.0  0.0  0.0  9.0  0.0
0.0  0.0  0.0  2.0  0.0  3.0  3.0  0.0  11.0
```

Fig 7

```
8.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
0.0  4.0  0.0  0.0  0.0  0.0  1.0  0.0  0.0
0.0  0.0  12.0 0.0  1.0  0.0  1.0  0.0  0.0
0.0  0.0  0.0  7.0  0.0  0.0  0.0  0.0  0.0
1.0  0.0  1.0  1.0  7.0  1.0  0.0  0.0  0.0
0.0  0.0  0.0  1.0  0.0  4.0  0.0  0.0  0.0
0.0  0.0  0.0  0.0  0.0  1.0  8.0  0.0  0.0
0.0  0.0  0.0  1.0  0.0  0.0  0.0  11.0 0.0
0.0  0.0  0.0  1.0  0.0  1.0  0.0  1.0  12.0
```

Fig 8

Fig 7 represents the confusion matrix when we used machine learning techniques. And fig 8 represents confusion matrix when we used deep learning algorithms.

	Method 1	**Method 2**	**Method 3**	
Technology used	Clarifai API	Spark	Tensor flow	
Languages Used	Java	Scala	Python	
OS Used	Windows	Ubuntu	Ubuntu	
Implementation	Easy	Hard	Very hard	
Accuracy	70%	68%	98%	
Running time	Low	Medium	High	

Results	Good	Good	More accurate	

6.Discussion & Limitations:

From the given data we can say that, we are able to analyze only for few historical images. Although there are so many places and statues, we couldn't find the find the data set for those images. The data set images i.e. the training data is not sufficient to analyze all the images which the user may upload now.

The main limitation which we got was collecting the training data set. We want to extend this into a mobile application too. Here one of the advantage of having the mobile application is it is easy to find the persons location easily when compared to know the location of a system when a person is using the web application. So we want to extend this to mobile application too.

Although it is useful, it is not easy to integrate the normal mobile application with the machine learning and tensor flow.

There are few other things which needs to be considered. Although we have started project by keeping in mind that whatever might be the image which a user uploads we should be able to get the details of that image. It might be a street, a hospital or a government building etc. But to achieve this we need to have huge amount of data sets and algorithms which can access these images fast and give us the output. But we couldn't find the dataset to implement everything which we thought.

And one of the other limitation is we tried to implement video analysis and give the description about the place. When we tried this we were able to get the information from the videos i.e. extract the information from the videos but we were unable to match those things with the other data. For suppose if we get data like building, its color, sea etc we couldn't be able to get the result of what we expected.

7. Conclusion:

From the above given description and implementation, we can see that we have succeeding in analyzing the historical monuments of so many images. we need to extend it to analyze statues of some of the world's most beautiful statues like statue of liberty. This further project work needs more set of data sets and modified algorithms so that we can extend this to video analyzing of places too.

8. References:

[1] https://github.com/nteetor/Image-Classifier
[2]http://www.di.ens.fr/willow/teaching/recvis09/final_project/
[3]http://cs229.stanford.edu/proj2013/KrzesinskiWilder-ImageObjectClassification.pdf
[4]https://sites.ualberta.ca/~bang3/files/DogCat_report.pdf
[5]https://github.com/KarthikVundela/BigDataAnalyticsandAppsCS5542_Project