Indian Institute of Technology Ropar

Department of Computer Science and Engineering

CS507: Course Project Report

Smart Album Creator

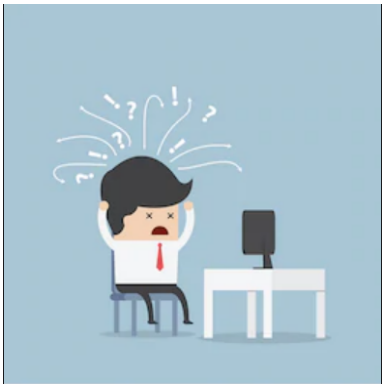
Authors
Piyush Pilaniya (2016csb1049)
Prinshu Kumar (2016csb1051)

Supervisors
Dr. Mukesh Saini

November 19, 2019

Introduction

Suppose you have thousands of photos in your laptop in which the photos are having various location(Mountain, beach, home) and themes(you with your friends, you enjoying in your college cultural fest). Now you want to see what varieties of photo you have in your album and see the best photos in those different varieties. Here comes the use of our smart album Maker which will do the job for you.



Objectives

Given a set of images in a folder, finding the images having various themes that vary in their nature and present it to the user in a directory having those photos. The best picture in each theme will be presented
So basically we will be using two medias:

- Images : Finding the description of the image in the textual form.
- Text : Using the description of the image and making the clusters of those images.

Implementation

For the implementation part we have taken following steps:

- Take input is a folder having a set of images.
- Finding the description of the images in the textual form.
- Clustering the images using their description and then finding the best quality photos in each cluster.
- Output is the folder having the images with various themes where every photo is the best photo of its corresponding theme.

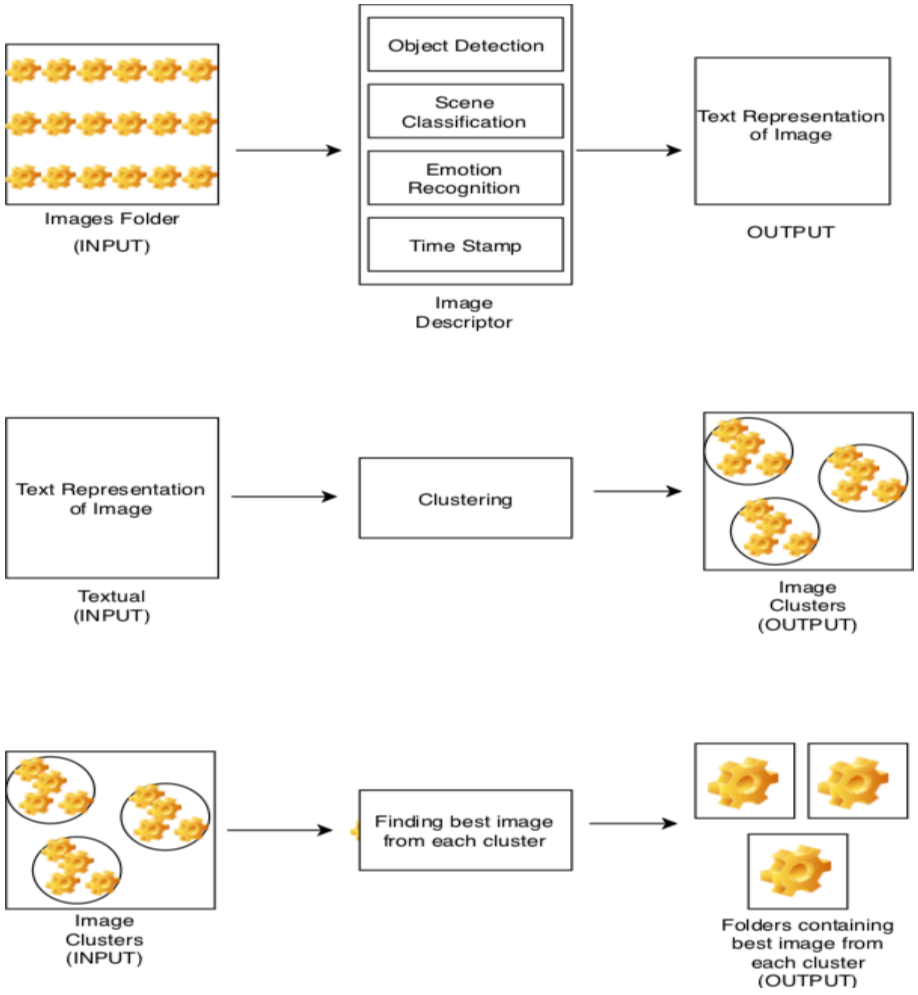


Figure 0.1: Flow diagram of model

Working of the model

The project is basically divided into three tasks:

Task 1: To extract the textual features from the image:

- Object Detection : resnet50 is used to detect objects in the image.
- Time Stamp : Pillow is used to extract the time stamp info from images
- Scene Classification : resnet100 is trained on the dataset of scenes taken from Kaggle competition held by Intel (https://www.kaggle.com/puneet6060/intel-image-classification)

Figure 0.2: Object Detection : Showing two object - person and tie

- Emotion Recognition : The dataset used is taken from (https://www.kaggle.com/c/3364/download-all). A much needed help was taken from (https://github.com/oarriaga/face$_c lassification$).

Task 2: Clustering Agglomerative Clustering which is also known as bottom-up approach is done in which initially each data point is a singleton cluster. Time stamp played an important role in our clustering. We have done initial clustering based upon time stamp. Images which fall in 120 seconds time gap are clustered together since they have high chance to fall in same theme and context of image. The text information extracted in task 1 is used here.

Task 3: Picking up best image from cluster. After all the cluster are formed which contain images of same person in same environment clicked at almost same time with expressing similar emotion, the task is to find which image is best from this cluster. For that purpose following things have differentiated images from being worst to best:

- Blur Detection : The variance of Laplacian is calculated to find the blurring in the image. Consider the below two images:
- Closed Eye Detection : A very common way to tell if image is good or not is to look at the eyes of all the person in image, if a single person exist with closed eyes then the image is not good. OpenCV is used to detect closed eyes in the image.

The image in which all the person have open eyes and the one with least blurring is preferred to be the final output.

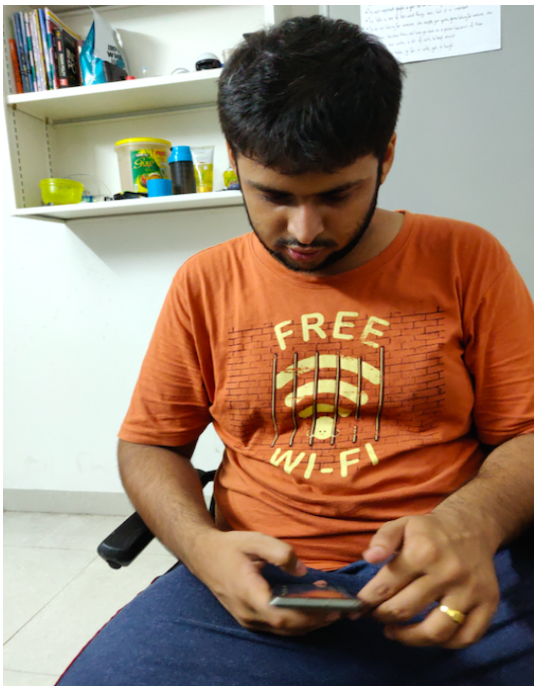


Figure 0.3: Blur value : 11

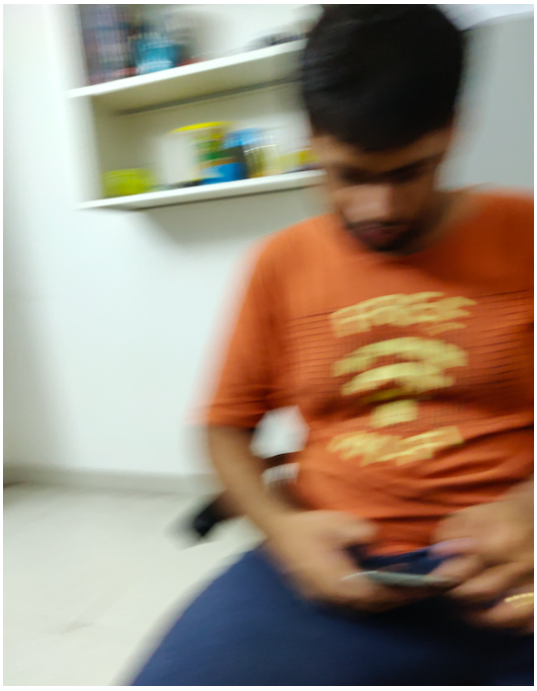


Figure 0.4: Blur value : 215

Results

We have obtained quite promising results. The output folder contained those best images which any normal human would select from his/her gallery. To verify the result we have asked various volunteers to create an output album manually and most of the images overlapped. For initial testing we have used 15 sample images with 6 images of environment which does not contain any person and other images are from phone gallery. After processing, the output folder contained 4 images.

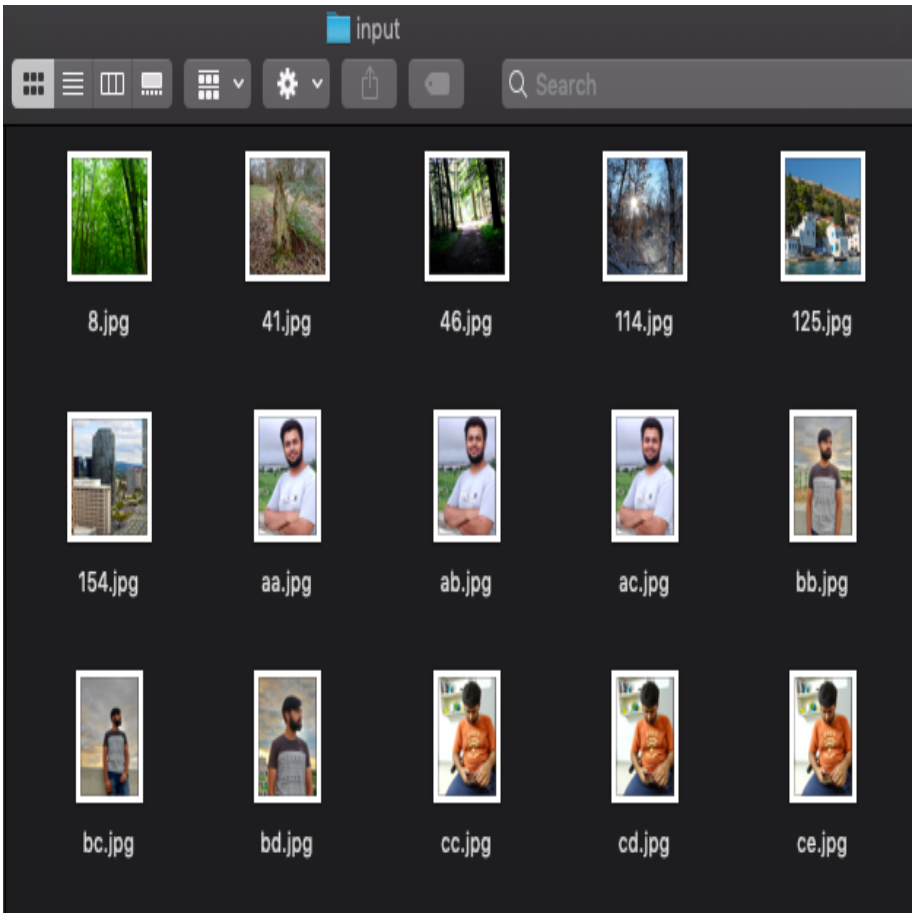


Figure 0.5: Input folder

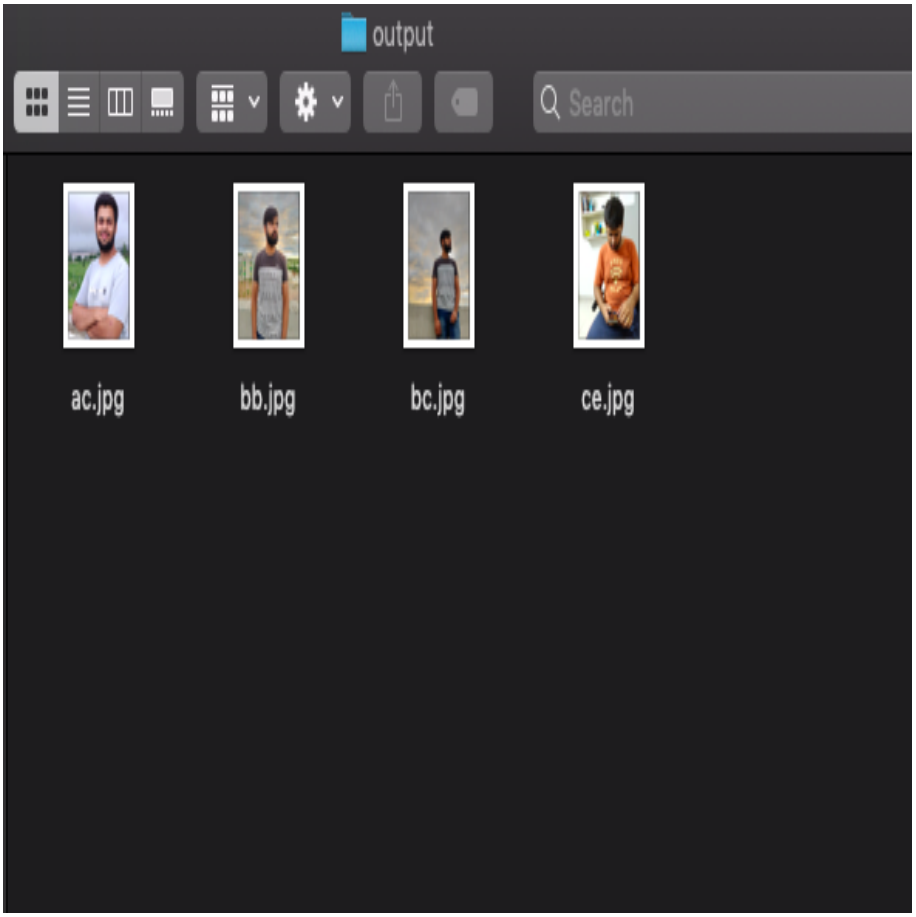


Figure 0.6: Output folder

Conclusion

We have observed that time stamp have been an important measuring factor in out model. We have done a series of experiments to get the optimum threshold value for agglomerative clustering, blur detection and image similarity measurement. we have also given the user the flexibility to put in a threshold value depending upon their need.

References

https://www.kaggle.com/puneet6060/intel-image-classification
https://github.com/NishantBhavsar/intel-scene-classification
https://github.com/omar178/Emotion-recognition
https://www.geeksforgeeks.org/ml-hierarchical-clustering-agglomerative-and-divisive-clustering/

https://www.pyimagesearch.com/2015/09/07/blur-detection-with-opencv/
https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/