Name: Trần Đình Tú (1651036)
Đinh Lê Quốc Việt (1651038)
Lê Duy Chương(1651060)

FINAL PROJECT REPORT

I/ Introduction:

My topic is about Image retrieval which using 845 images from Oxford Building Dataset. Our input is an image and it will return top 5 similar images. For each image and landmark in our dataset, one of four possible labels was generated:

1. *Good* - A nice, clear picture of the object/building.
2. *OK* - More than 25% of the object is clearly visible.
3. *Bad* - The object is not present.
4. *Junk* - Less than 25% of the object is visible, or there are very high levels of occlusion or distortion.

I/ Approach:

Our approach has 4 steps:

_ Firstly, feature extraction and description uses Harris Laplace corner detector and SIFT descriptor. We choose this feature because our dataset is about buildings which have many corners and we can use this feature to distinguish between buildings. We also apply laplace or DOG(Difference of Gaussian) to help the feature to be invariant to scale. After that, we describe each key points by an array of length 128 by SIFT algorithm.

_ Secondly, clustering uses KMean. We will clustering the key points into different clusters in order to gather the key points that are similar to each other. In additional, it is also easy to represent the image.

_ Thirdly, image representation uses TF-IDF. We represent our image by an array that has length equaled the number of clusters. Each value of the array is the TF-IDF or weight of different clusters in the image. With TF-IDF, we can figure out with type of key points is important in the image and which ones are not.

_ Fourthly, similarity calculation uses Cosine Similarity. In this step, the similarity between the input image and the image in the data set is calculated by the Cosine Similarity of the array that represents our image. Then, we will choose the images that are most similar to our image.

III/Experiment:

_ We implement our approach in Colab environment using Python 3, OpenCV, sklearn, numpy libraries

_ In the first step, we extract key point by Harris Laplace corner detector in OpenCV. The max number of corner key points in the image is set 1000. We chose 1000 because we see that the average key points is approximately 1000 and we consider 1000 is enough to represent an image.

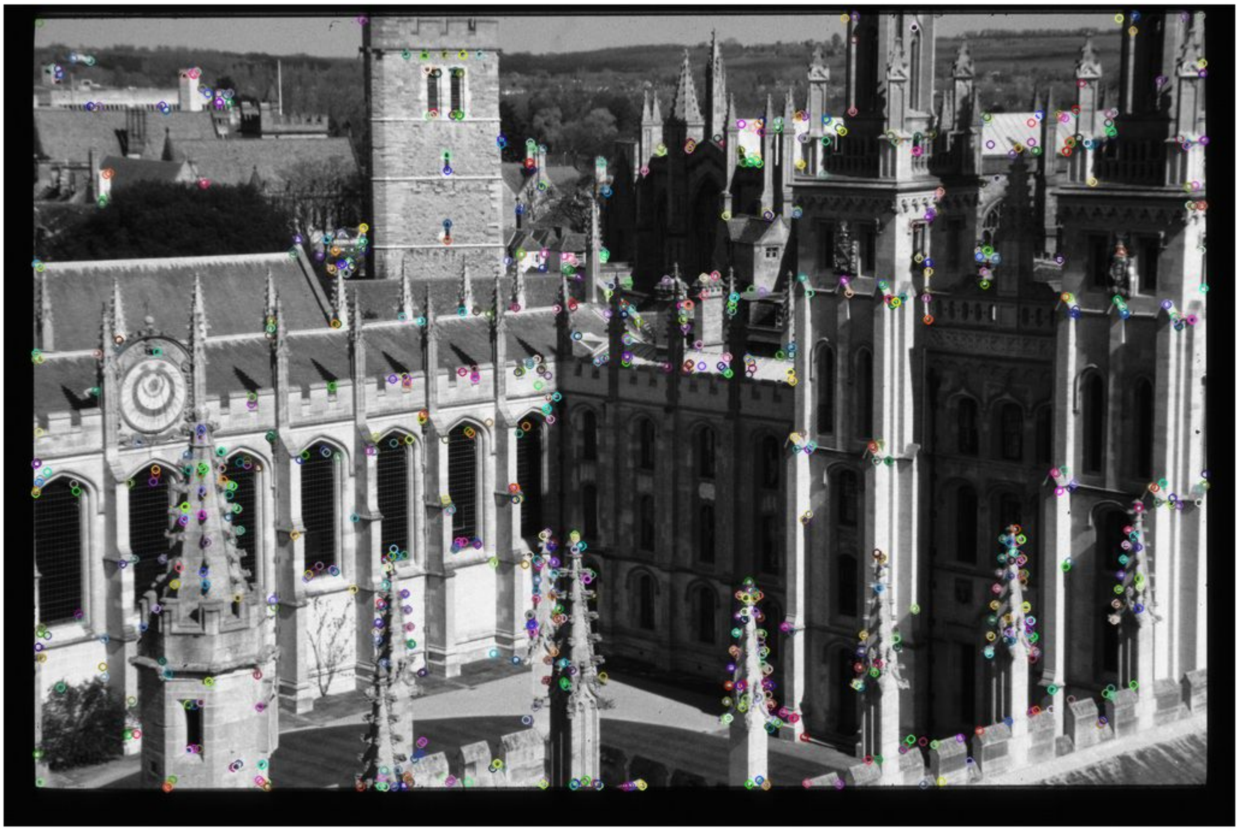


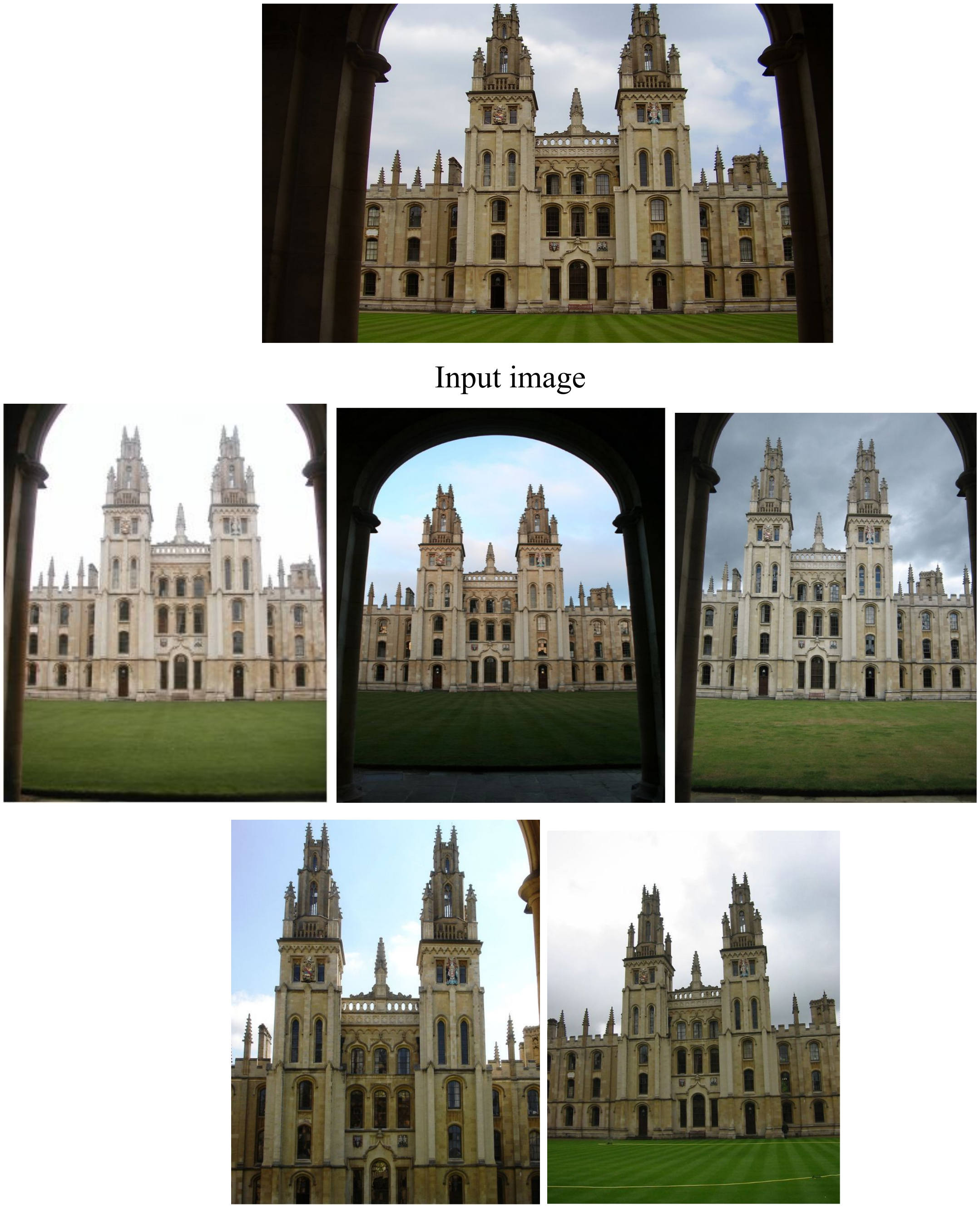
Figure 1. Image with corner key points.

_ In the second step, we chose the number of clusters is 1000. This number was chosen because the number of maximum key points is 1000 and we want to compute faster. The KMean is implemented by sklearn.

_ In the third and fourth step, we just calculated as usual and the Cosine Similarity is also uses sklearn library.

_ According to 3 types of image in dataset, we have the result here:

+ Good type

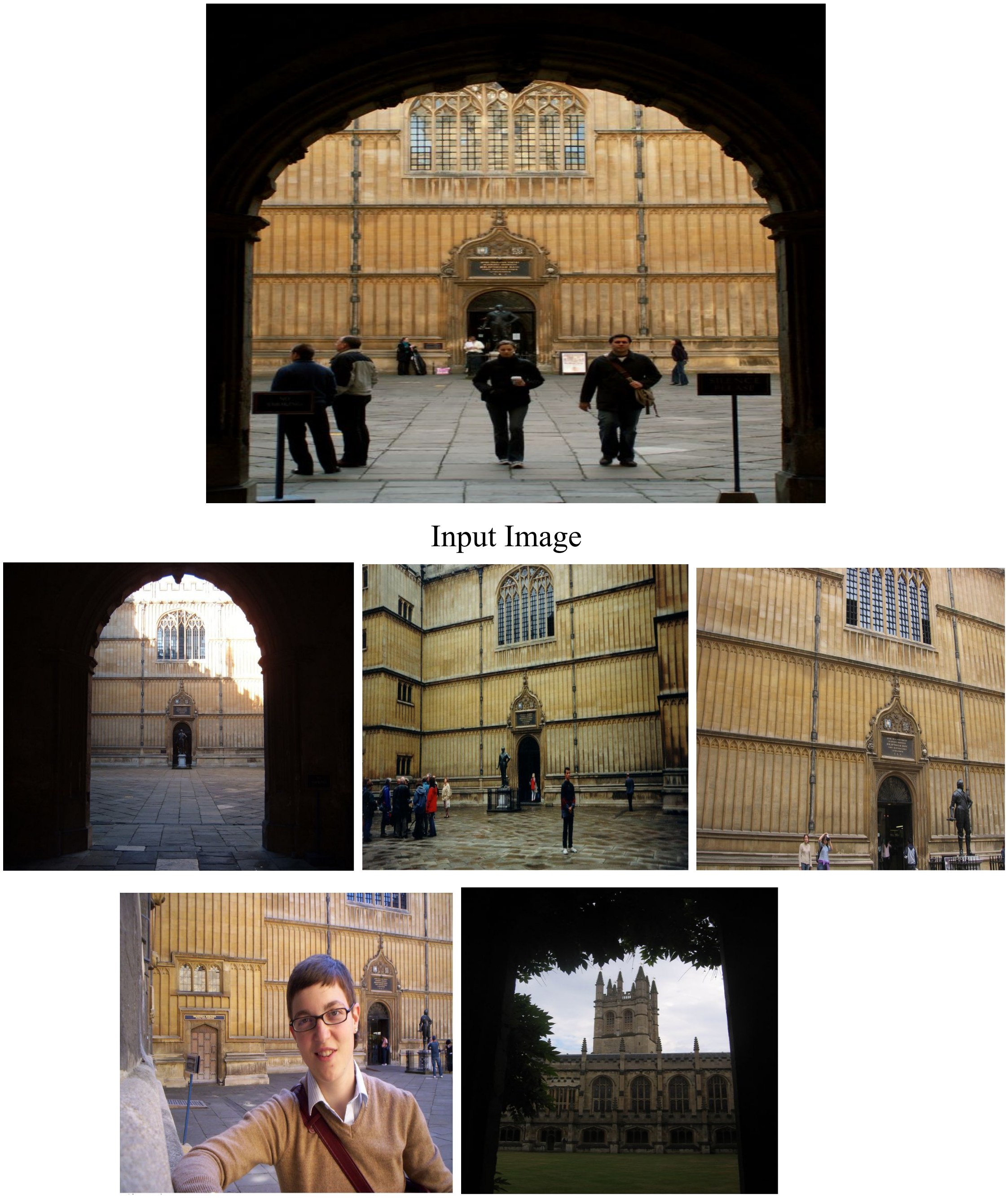


Input image

. Top 5 Output Images

Figure 2.

+Ok type:

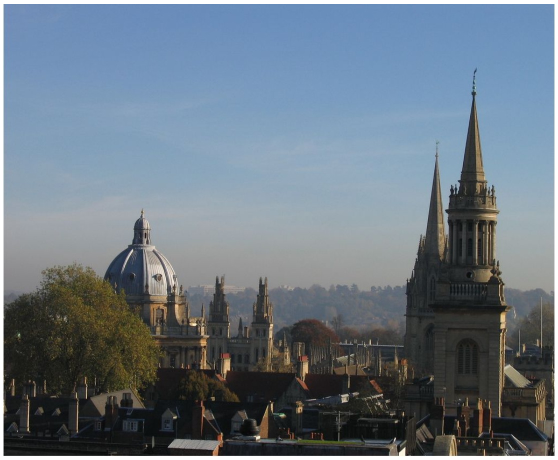


Input Image

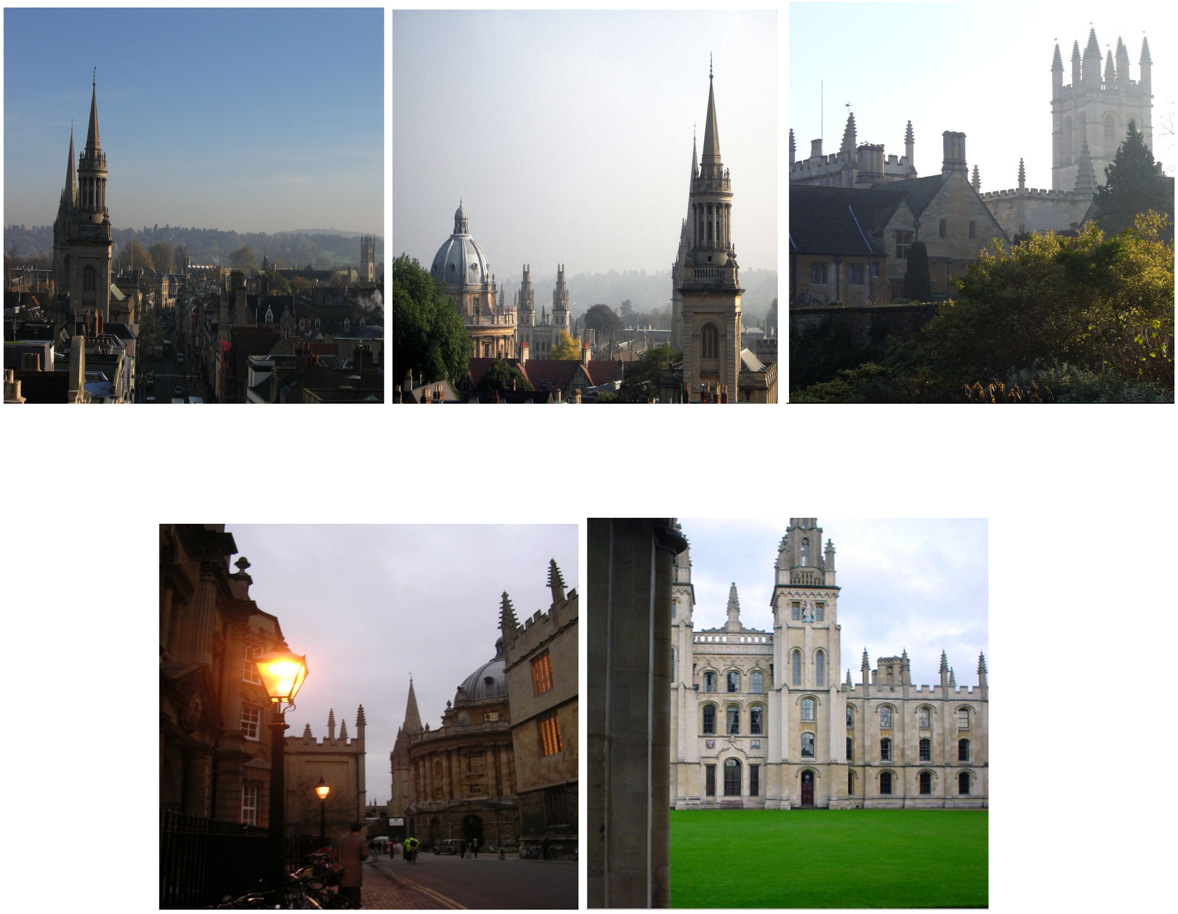
Top 5 Output Images

Figure 3.

+Junk type:



Input Image



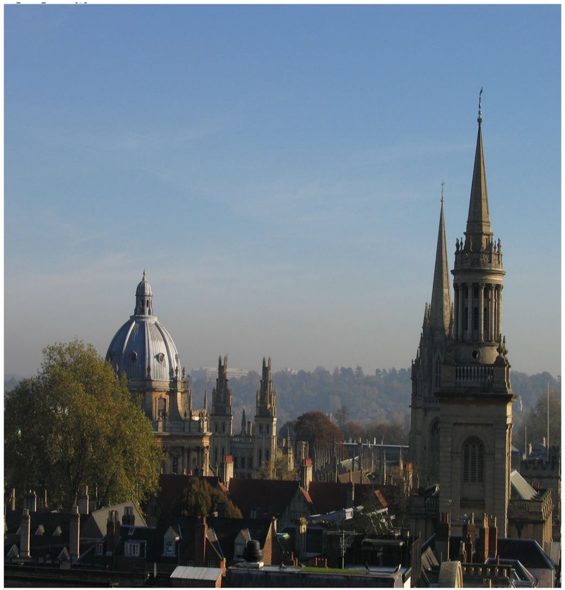
Top 5 Output Images

Figure 4.

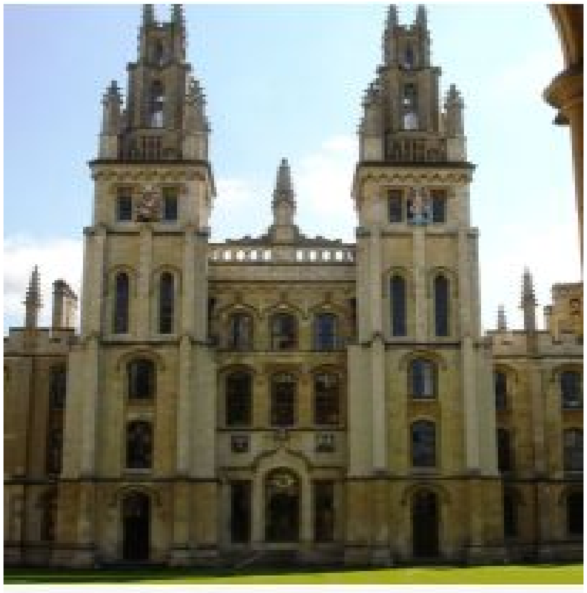
_ Because the evaluation in the dataset is not clear so we just evaluate by running some pictures and see the result like this. So we see that with the image in type “Good”, we have a really good result. With the image in type “OK”, most of the output images are related to the query except the last one. With type “Junk”, the result gets worse while just 2 images are related and the other seem to be not. So we figure out that with the image that has clear picture of object we will get better result.

IV/Conclusion:

_ The result is really good but it still has drawbacks. Firstly, the computation of clustering is still slow, it takes a lot of time. Secondly, the feature does not contain spatial information so sometimes it retrieved the images that has the building not in the same region like figure 5. Thirdly, the problem of sparse vector of image representation. So in the future we will try to add spatial information by triangular regions based technique and also try to solve other problems



Input Image



Retrieved Image

Figure 5