CPSC 425: Assignment 2

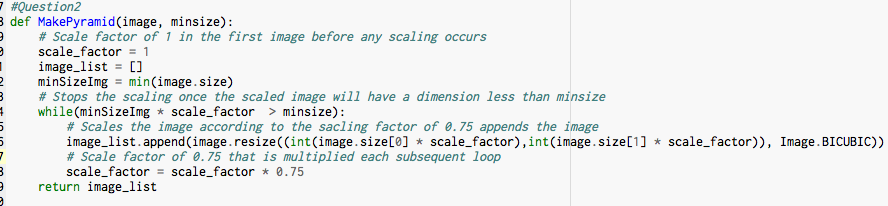
Name: Terence Chen

Student number: 42602136

Q1)

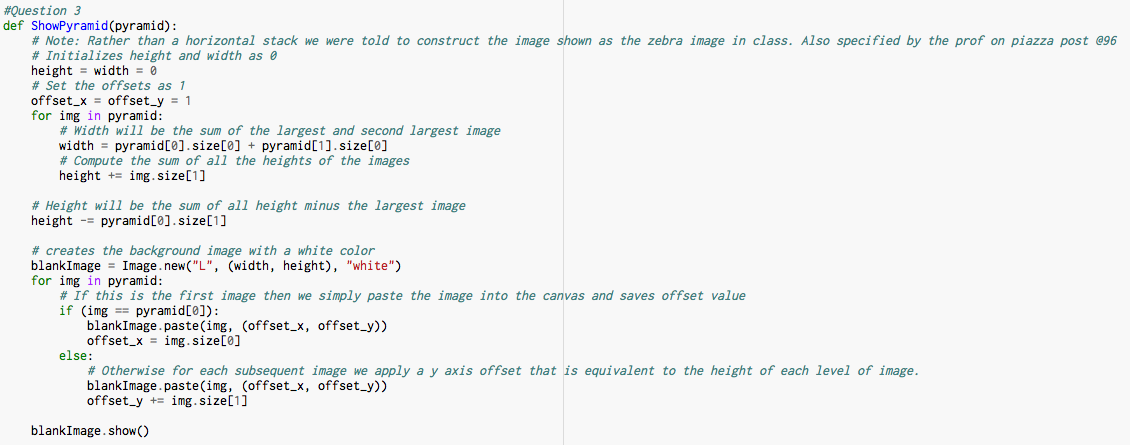
I unzipped the files as specified in the assignment description

Q2)

Code: 

Q3)

Code:



Using a min size of 50. ShowPyramid(MakePyramid(sports, 50)) I get the image: 

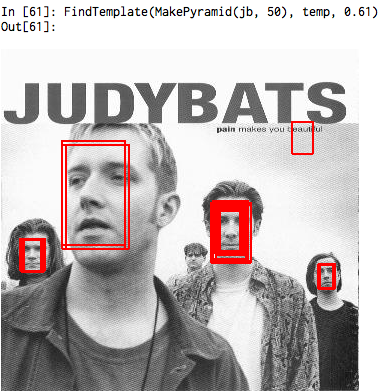
This stacking layout was specified by the prof in class and discussed in piazza, on post 96.

Q4)

Code: 

Q5) So for question 5 I simple tried out different threshold values then calculated the error rate by counting the false negatives and false positives by manually checking each image. I notice that as we get to the range of 0.55 – 0.65 we get pretty good results. A relatively good threshold value is 0.61. I will show the images below using 0.61 as the threshold value.

judybats:



students:



tree:



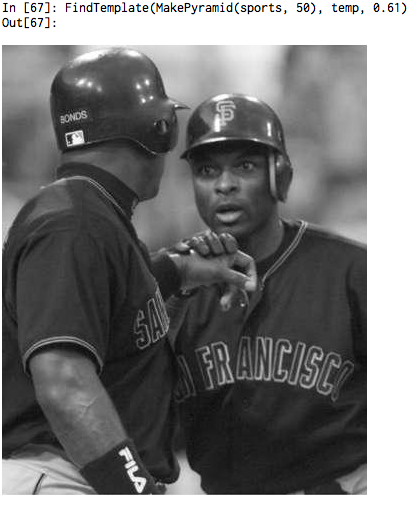
family:



fans:



sports:



Q6) A recall rate is defined as the set of retrieved and relevant documents over the set of all relevant documents. I am going to assume that every face on the image is a relevant document. Using the above images I get the following recall rates.

For “judybats” we have a total of 4 faces and 3 retrieved thus we get a recall rate of 3/4.

For “students” we have a total of 27 faces and 22 retrieved thus we get a recall rate of 22/27.

For “tree” we have a total of 0 faces and 0 retrieved thus we get a recall rate of 0/0 (undefined).

For “family” we have a total of 3 faces and 3 retrieved thus we get a recall rate of 1.

For “fans” we have a total of 3 faces and 0 retrieved thus we get a recall rate of 0/3.

For “sports” we have a total of 1 faces and 0 retrieved thus we get a recall rate of 0/1.

We see that we get very low recall rates for certain images, I think this is because of how we defined what belongs into the relevant document set. The NCC finds matching based on the similarity between our template image and the corresponding images. Since I defined that all faces belong in the relevant documents, for images like the sports.jpg, tree.jpg and fans.jpg there are not a lot of similarities between them and the template image we used. While for images like judybats.jpg, students.jpg and family.jpg the faces in these images are a lot more similar to our template image thus we get more correct matchings which increases our recall rate.