Assignment 1

The dataset used in Assignment 1 is a subset of NUS-WIDE. It consists of 25 object categories and each category has 50 images as database and 10 as queries. For more details about NUS-WIDE, you may refer to the attached paper named **NUS-WIDE: A Real World Web Image Database from National Univeristy of Singapore.pdf** and this link http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm

The information about each folder is illustrated as follows:

- The names of the 25 object categories are stored in **.\ImageData\category_names.txt**.

- Database for image retrieval are stored in **.\ImageData\train**
  - Images are stored in **.\ImageData\train\data**, each class has a folder consists of 50 images.
    - Some images may be contained in more than one category. For example, if one image contains a bear as well as trees, it may appear in both the folder named bear and the folder named tree.
  - The ground truths of images are stored in **.\Groundtruth\train**.
  - Some tags of the images are stored in **.\ImageData\train\train_tags.txt**.
    - For image tags, each line in the ***_tags.txt** consists of image_id and its tags.

- Database for query are stored in **.\ImageData\test**
  - Images are stored in **.\ImageData\test\data**, each class has a folder consists of 10 images.
  - The ground truths of images are stored in **.\Groundtruth\test**.
  - Some tags of the images are stored in **.\ImageData\test\test_tags.txt**.

- The ground truth for the database is stored in **.\Groundtruth\train** (.\Groundtruth\test).
  - Each object category has a corresponding "Labels_categoryname.txt" file. In this file, the order of images is the same as the order of images in **.\ImageList\train\TrainImagelist.txt** (.\ImageList\test\TestImagelist.txt). Some images may be contained in more than one category.

- A working image search system (including UI) that uses only color histogram feature is stored in **.\ImageSearch**. The main function is in ImageSearch.java. You need to incorporate and combine other features.

- **.\ FeatureExtractor** folder consists of 4 folders to introduce the following features.
  - **colorhistgram** : Please refer to the **readme.docx** in the **.\ FeatureExtractor \ colorhistgram** for more detailed information.
  - **Semantic Feature :** Please refer to the **readme.docx** in the **.\ FeatureExtractor \ semanticFeature** for more detailed information.
  - **SIFT + Sparse Coding features**: Please refer to the **readme.docx** in the **.\ FeatureExtractor \ SIFT+SC** for more detailed information.
  - **SIFT feature demo**: Please refer to the **README** in the **.\ FeatureExtractor \ siftDemoV4** for more detailed information.

If you have any question on the dataset, please email to me. jingyuanchen91@gmail.com