Cognitive Computing – Fall 2019

HW#2 (DUE: 10pm, Saturday, 16 November, 2019)

Note:

1) If you have any questions regarding the homework, send e-mail to the TA at <d06944009@ntu.edu.tw>.
2) Submit a soft copy of your results and codes (in *Jupyter Notebook*, see explained below) to TA **before the due** via CEIBA. Name your zip file with your student ID. i.e. d06944009.zip. The images are not required.
3) The dataset is available at: http://bit.ly/2Wmqi2t
4) You are highly encouraged to write the homework in English.
5) Please DO write appropriate comments along with your codes.

The goal of the homework is to *design* and *evaluate* (and the sense) the visual features across different image categories measured by image matching (or retrieval) performance. For this work, we will use an e-commerce dataset containing 30 categories, where each has 20 images. Though having minor issues, we treat each category label as the (search) ground truth. You can go through the dataset for more information.

Also remember to visualize your query results (at your own and NOT needed to be submitted) and see if it matches the performance evaluation.

You will evaluate with **color, texture/shape, and local features** for the homework, and decide the proper similarity or distance metrics (e.g., L1, L2 distance, cosine similarity, etc.). Please brief the features and the similarity (distance) metric you adopt.

1) *Color similarity*. You can refer to the color features introduced in the lectures, or other relevant papers. Suggestions for color features are (but not limited to): *global color histogram, regional color histogram, grid color moments, means in each color channel, color (auto-) correlogram*. You are encouraged to implement on your own or can find other open source tools; however, you have to acknowledge the source of the tools.
2) *Texture/shape similarity*. You can choose any texture or shape features. Suggestions for such features are (but not limited to): *Fourier features, Laws' texture measures, co-occurrence matrix metrics, Tamura's textures, Gabor texture, PHOG, gradient histogram, edge histogram, etc*. You are encouraged to implement on your own or can find other open source tools for texture or shape features; however, you have to acknowledge the source of the tools.
3) *Local feature similarity*. We suggest DoG + SIFT for implementing local features. You are encouraged to use the <u>visual word representations</u> (as introduced in the lecture) or some advanced methods for pooling local features (e.g., Fisher, VLAD, etc.). It will incur some technical issues when doing clustering with large codebook size (e.g., K = 100k). Will suggest using smaller codebook size. The provided dataset is for experimenting for ranking (testing). You can use other datasets to train your code books or use others' code books. You should be able to find some local feature code books available. You are OK to use others but do cite the source properly.
4) **Evaluations**. We are keen to different modalities and the performance. Here we are to evaluate MAP (mean average precision across the whole dataset). Note that when you do the query, please

use the ALL reference images (~600) across categories as the candidates for ranking. We do the evaluation in leave-one-out fashion. Thus, every dataset image is used for evaluation; we can derive its own average precision (AP) with each setup. We can then use the 600 per-image AP to calculate the MAP across the dataset or the MAP per category (regarding to the feature setup). You also need to report the best 2 categories and the worst 2 categories for the category MAP. You can try to fill in the table below for the submission. Note that the fusion part is optional. You do not need to stick with the table format below but to aggregate the required information to compare. We are to list here for comparison the capabilities for different modalities.

categories vs. methods	MAP	Best Two Categories (w/ MAP)	Worst Two categories (w/ MAP)
Color	0.000	A (0.000), B (0.000)	C (0.000), D (0.000)
Texture (or Edge)	0.000	A (0.000), B (0.000)	C (0.000), D (0.000)
Local Feature	0.000	A (0.000), B (0.000)	C (0.000), D (0.000)
Fusion *(Optional)*	0.000	A (0.000), B (0.000)	C (0.000), D (0.000)

5) The homework is *NOT* entirely graded by the performance (though might considered). Try your best to see what you can achieve. However, we will invite few students having a short sharing for what you been doing for the work.

6) When doing the experiments, will suggest extracting the image features *offline* and save them as files first. It will speed up the experiments. You do not need to send the feature files.

7) Collaboration policy: It's OK to share your feature extractors with others. Or ask the help from others. We have the bulletin board (討論群組) open in CEIBA where you can interact with, share your implementations (in GitHub), or work out the HW discussing with others. However, **you need to do feature extraction, ranking, visualization, and evaluations on your own**.

8) For submission, please submit your code and a *Jupyter Notebook* hw2.ipynb to CEIBA, name your .zip file with your student ID. (.zip files only, please do not submit rar or 7z files).

For example:

d06944009/

----hw2.ipynb

----code.py

----data/

--------(unzipped data here)