FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO
MIEIC - 2015/2016
COMPUTER VISION

PROJECT Nº 2
Image classification

Summary

The main goal of this project is to develop a system that can classify an image as being a cat or a dog. This project is based on a past Kaggle competition available at https://www.kaggle.com/c/dogs-vs-cats.

Image classification system

There is no restriction on how the system should be built. Nevertheless, a typical image classification system includes the following modules:
- **Feature detection**: A local interest point detector to find relevant features.
- **Feature description**: A local descriptor scheme to extract invariant descriptors.
- **Image representation**: All the descriptors extracted from all the images are clustered (e.g. k-means) to quantize the space into "visual words". Each cluster mean represents a word of the resulting dictionary. An image is then represented by a histogram of words, using the vocabulary built previously. The resulting bag-of-words representation vector should be normalized.
- **Classification**: A classifier is trained with the training images defined for the dataset. After having the model trained, the input image is tested in order to obtain a category.

Additional modules can be implemented and will be evaluated as extras:
- **Evaluation**: Compare at least two different approaches, for example, different classifiers or different local descriptors studied in the Computer Vision course.
- **Colour descriptors:** Local descriptors do not use the colour information. However, this information can be useful in a system like this. Combine colour descriptors (e.g. dominant colour, histogram of colours, etc) with the bag of words representation.
- Other improvements will be considered in the evaluation, if justified.

Scientific Paper and Delivery

A short report must be elaborated in the format of a scientific paper (max. 4 pages), including:
- Brief introduction to the problem, including references about the state of the art;
- Description of the developed system;
- Possible additional specifications or improvements;
- Results of the image retrieval system, namely percentage of categories correctly identified and other measures considered relevant (e.g: true positive rate, true negative rate, accuracy, etc.);
- Discussion about the overall performance of the system and possible situations where it fails;
- (extra) Comparison of performance using different approaches;
- Conclusions and future improvements.

The paper can be written in English or Portuguese and should be based on the model available in Moodle. The code, with meaningful comments, should be presented in annex.

The work must be submitted at the Computer Vision page, in the UP Moodle site, until the end of the day December 14, 2015.

Bibliography and other support material

- Visual Categorization with Bags of Keypoints, C. Dance, J. Willamowski, L. Fan, C. Bray, and G. Csurka, ECCV International Workshop on Statistical Learning in Computer Vision, 2004.
- Sampling Strategies for Bag-of-Features Image Classification. E. Nowak, F. Jurie, and B. Triggs. ECCV 2006.
- Video Google: A Text Retrieval Approach to Object Matching in Videos, J. Sivic and A. Zisserman, ICCV 2003.
- Object Recognition from Local Scale-Invariant Features, D. Lowe, ICCV 1999.