COMPUTER VISION

ASSIGNMENT 4

Assoz.Prof. Antonio J. Rodriguez-Sanchez, 17-June-2023

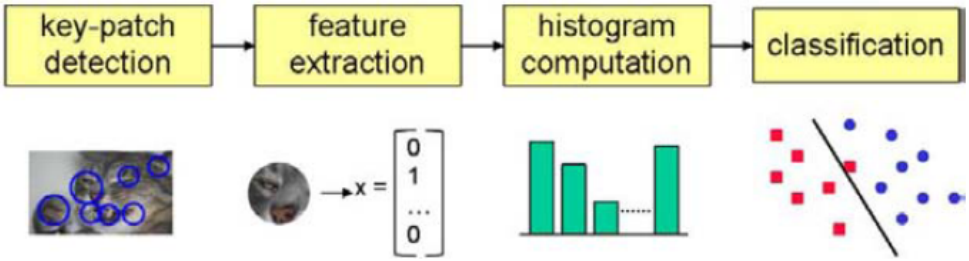
Deadline: 07-July-2023

This assignment can be done in groups of up to two people

(15 points) Construct a basic object recognition system (bag of words) and compare to a simple Convolutional Neural Network (CNN).

1. GOAL: Develop a system that can detect leafs or aircraft signals (see below) present in images. The object recognition system will be given an image and its output is to provide if contains the object to be recognized.

    a. The first system will follow the process explained in lecture 10 (slide 66), namely:



    b. The second system will make use of Neural Network, namely a CNN.

2. BAG OF FEATURES.

    The Features:
    You will use three features:
    a. Shi-Tomasi corner detector.
    b. HoG (Histogram of Gradients).

    These are popular feature descriptors available online and in OpenCV.

    The classifier: Use a Support Vector Machine (SVM). Look for a library (one is given at the end of this document) and learn how to use it.

3. THE DATASET AND DATA PREPARATION:

    a. Choose one task from:
        - Leaf dataset: https://data.mendeley.com/datasets/hb74ynkjcn/1
        - Aircraft signal dataset: https://github.com/yalesong/natops

b. Read the images and assign their corresponding labels.
   - Consider that the images are at different sizes, you must transform all images to the same size.
   - Colored pictures may have to be transformed to black and white (bag of features).
   - Usually images are not normalized regarding contrast, if needed this can be done easily with just one line of code such as (this is Matlab):

```
img2 = uint8((double(img)-double(min(min(img))))*(255/
double(max(max(img))-min(min(img)))));
```

4. BAG OF FEATURES. OBJECT REPRESENTATION: feature extraction using Shi-Tomasi/HoG.

   a. After the keypoint localization, you have to convert them into their feature vectors.
   b. Those feature vectors must be clustered into a histogram, for this the most used method is k-means (available in OpenCV: MiniBatchKMeans).

	Feat. 1	Feat. 2	Feat. 3	Feat. 4	Feat. 5	Feat. 6	Feat. 7	.....	Feat. M
Image 1									
Image 2									
Image 3									
.....									
Image N									

5. LEARNING: Training a classifier. Download the code for a classifier. You will use an SVM (e.g. library libsvm for Python or any other one). For comparison you will adapt the code of CNN explained in class, but replace VGG for Resnet-18. The CNN must also be adapted to work with your dataset. For the CNN you do not need to create a histogram, just feed the images to the CNN and record the accuracy and confusion matrix. Remember to use data augmentation.

6. RECOGNITION: Using classifier in test data. We have now a trained system that can distinguish an image (leaf type, aircraft sign).

7. RESULTS

   At this point we have a classification system that can recognize what is on the picture. We have to measure how good is our system, for this we compare the true labels of the test set with the ones obtained in step 5. The higher the percentage the better, if you have 2 classes and this percentage is below 50% it means our system is ever worse than using chance (a coin with tails and heads), if it is a bit over 50% it means it is ok, 70-80% good enough and over 90% quite good, if over 95% it is great. Other cases: if you have 10 classes, you should get over 10%.

Create first a confusion matrix, from this, you can build a table reporting true positives, false positives, true negatives and false negatives.

Compare the three features for the two classifiers (Bag of words with SVM and CNN): Shi-tomasi, HoG to see which one performs better at the task. This should be reported in graphs, tables or ROC curves.

Deliverable: Document with the following:

Page 1

a) Introduction on object recognition, the steps to achieve object recognition.
b) The purpose of this assignment. State the *hypothesis*.

Page 2

Your solution: How your solution was implemented and the algorithm you followed to solve the assignment. How the features were created and what are their basis for their creation. Comment on the pattern recognition approach: SVM and CNN.

Page 3-5

Results: Include the results of the assignment. Provide cases where the method worked and cases where it did provide the wrong answer. Give the success rate (percentage of correctly classified images) and TP/TN/FP/FN. Provide graphs, confusion matrix and tables.
Compare the results from the Shi-tomasi and HoG for each classifier.

Page 6

Discussion: What you learnt about this assignment and how the method can be improved or extended.

Appendix

Code and anything else you may consider of interest

Additional material

An example on Bag of Words with OpenCV:
https://www.kaggle.com/pierre54/bag-of-words-model-with-sift-descriptors

For the CNN, use Resnet-18, that is, not the VGG as the latter will take too much computation time to compute.

HoG: http://rogerioferis.com/VisualRecognitionAndSearch2014/material/papers/DPMCVPR98.pdf

Libsvm: https://www.csie.ntu.edu.tw/~cjlin/libsvm/