INF6804 Computer Vision
H2024 – Practical Assignment #1
Description and comparison of regions of interest

Objectives:

- Introduce the processing of video and image data using computer vision libraries.
- Teach about image description based on Co-occurrence Matrices (CM).
- Teach about image description based on Local Binary Patterns (LBP).
- Learn to compare features vectors.
- Explain the advantages and shortcomings of both image(s) description approaches.

Submission:

- *All your source code* (we should be able to run your tests)
- A report (*.pdf format* of 8 to 12 pages with font size of 10)
- Submit before Feb. 9th 2024, 5:00 PM, on Moodle –*late submissions will not be accepted*
- *You must also submit your report on TurnItIn*
  - Register at www.turnitin.com using the info available on Moodle!

References:

- See course notes on Moodle (Chapter 2)

Other directives:

- The assignments must be made in teams of two, submit only one version of your work!

Presentation

In this assignment, you will characterize two methods used for the description of regions of interest in images: **Co-occurrence Matrices (CM)** and **Local Binary Patterns (LBP)**. You will determine which method performs better under various circumstances, and your report must include a description of your work, experiments, and answers to outlined questions.

You have to compare two methods of description: the first method is based on **Co-occurrence Matrices (CM)**, the second one is based on **Local Binary Patterns (LBP)**. You can use your course notes as a reference to understand their basic working principles. For more details, go look online! To compare both methods, you will have to solve the task of **content-based image retrieval (CBIR)**. You will have to characterize each image using descriptors derived from GLCM and LBP and then find the images that correspond to each request image by measuring the similarity between these descriptors. We ask you to compare the two approaches in terms of efficiency (computation time) and performance. You must evaluate their performance based only on the result of **the 3 most similar images** that you have extracted from the database. You are free to choose the metrics that, based on your knowledge, fit the best for the methods to implement.

Using data provided in **"data.zip"**, you have to include the following in your report (marked on 20 pts):

1. Presentation of the two compared approaches (3 pts) :

   In your own words, give the general description and principles of your two approaches.

2. Performance hypotheses in specific use cases (2 pts) :

   Identify, based on your theoretical understanding of the two approaches, which one should be the best of the two in at least TWO specific use cases. For example, which is the best method to use if the size of the compared regions content is relatively uniform? Why?

3. Performance hypotheses concerning the bounding boxes (1 pts):

   Knowing that you have access to the bounding box of the object of interest in each images, Do you expect improvement in performance for both methods? Why?

4. Description of experiments, data and evaluation criteria (3 pts) :

   Describe in detail the experiments realized to test the hypotheses from point 2 and 3. What are the difficulties of the request images in the database? Which measures of similarity between descriptors have you chosen? Which evaluation metric did you use?

5. Description of the implementations used (2 pts) :

   Describe the implementation of the two studied approaches. If you did not write all the code yourself, where does it come from? Did it require modifications? Otherwise, from which papers or websites did you inspire yourself to write it? In all cases, what are the primary parameters of your methods? How did you set their values?

   Conduct a hyperparameter search to fine-tune 'P' (number of sampling points) and 'R' (radius) for LBP, 'levels' (quantization of levels) and choice of channels for CM, ensuring optimal descriptor performance.

   You can set the bounding boxes manually or compute them.

6. Experimentation results (3 pts) :

   Provide the evaluation results from your experiments related to the hypotheses of the second and third point. Use a proper format for their presentation — tables, figures, ...

7. Discussion on results and prior hypotheses (3 pts) :

   Discuss the results of the sixth point in relation with the hypotheses of the second and third point. Which hypotheses are supported by these results? Which are not? Which test resulted in a lack of conclusion? How could you improve these tests?

   Additionally, assess how CM and LBP performed on *'dolphin_query.png'* and consider if another descriptor covered in class might be more effective for this case. why?

8. Readability and completeness (3 pts) :

   In addition to the content, the format must be clean and complete.

During the lab periods, do not hesitate to ask questions — we can help you with any technical issue if you are working on Windows/Linux, or if you are coding in C/C++, Python or Matlab.

You will be **penalized by 50% of the total grade if you do not hand in your code**. Also, if your report is not submitted to **Turnitin**, it will not be graded. The order of presentation for the topics listed above does not matter, as long as they are all present.

Resources

Vision libraries:

- OpenCV (https://docs.opencv.org/4.0.0/d9/df8/tutorial_root.html)
- scikit-image (https://scikit-image.org/docs/stable/auto_examples/index.html)

Python:

- Guide (https://wiki.python.org/moin/BeginnersGuide/Programmers)
- NumPy (https://docs.scipy.org/doc/numpy/user/quickstart.html)

Matlab:

- Guide (http://www.mathworks.com/help/pdf_doc/matlab/getstart.pdf)
- Cheatsheet (http://web.mit.edu/18.06/www/Spring09/matlab-cheatsheet.pdf)