ARI3129 - Advanced Computer Vision for AI
Assignment

Deadline: 23rd January 2022
Maximum Marks: 100 marks.

Part 1 - Investigating topics in CV (40 marks)

Team size: Individual

Length: 5 Pages, including figures and references
Format: IEEE Conference Paper[1]

Rationale:

You are expected to survey two traditional techniques related to the topic that you choose and compare them to two modern techniques that would use any form of deep learning techniques. You are strongly encouraged to find off the shelf open-source implementation of these techniques and run them for comparative purposes. You may also choose to implement an algorithm of your choice. There is no strict expectation and you will be assessed on the level of reporting presented in the paper.

Marking Scheme:
- Literature Review (4 techniques: 2 traditional and 2 deep learning based) 20 marks
- Methodology (execution of 2 third party open source implementation) 10 marks
- Evaluation (comparison of the 2 implementations): 10 marks

Deliverables:
- Paper in PDF format
- Zip file with the implementation
    - Folder with Traditional approaches
        - Folder for each technique
    - Folder with Deep Learning approaches
        - Folder for each technique

List of Research topics:

- Visual Saliency
- People Counting and/or crowd analytics
- CV based sports analytics
- Facial Recognition
- Hand gesture recognition
- Pose estimation

---

[1] https://www.ieee.org/conferences/publishing/templates.html

Part 2 - Applied CV Task (40 marks + 20 marks)

Team size: Maximum of 4 persons. Equal contribution will be assumed.

In this part, you are expected to solve this practical CV problem using Python and OpenCV on a Jupyter Notebook. You are expected to document your work throughout all the tasks.

Task 1 - Acquiring Data

Due to updated security protocols, capturing of online video streams is now beyond the scope of this assignment. For this reason, a set of videos are provided in the link below for you to use within this assignment:

https://drive.google.com/drive/folders/1IhQekUjSX0M8MuBzRyM5xVN7Ri_rtRDs

This link is accessible via your University of Malta account.

Marks - NA

Task 2 - Background Detection

Segmenting the background from the foreground is a challenging task. When working with video data the task becomes somewhat less complex as moving objects can be detected and removed from the scene in order to find the background. In this task, you are required to create a script that is able to extract the background from a video. The script should extend to two different parameters and should work either for the video stream or for the saved video files.

Marks - 7

Required

Script(s) with the above specifications. Document the libraries and versions needed to run your script.

Task 3 - Object Detection

At this point in your assignment, your program should be able to differentiate between background and foreground. By using Deep Learning and other Object Detection techniques apply object detection to a video feed. You can use any deep learning model of your choice and also experiment with training it using your annotated data. You should conduct the following experiments and compare results;

1. Object detection on live stream (without pre-processing)

2. Object detection on background only
3. Object detection on the foreground only

*Record 30 seconds to 1 min sample footages at different time intervals (morning, noon, evening) (You need to have at least a dataset of 2 - 3 min in all)
* To annotate your dataset you can use the following web application
http://labelme.csail.mit.edu/Release3.0/
* Other annotation tools may be found in this link:
https://lionbridge.ai/articles/image-annotation-tools-for-computer-vision/

Marks - 8

Required

Script(s) with the above specifications. Document the libraries and versions needed to run your script.

Documentation comparing the different results

Task 4 - Compare Object Detection Techniques

Rerun the previous object detection process but use two different techniques. You should use at least two computer vision models (such as YOLOv3.0, Retinanet50, Mask R-CNN or any other of your choice) as well as one Computer Vision or Machine learning Model (Template Matching, HoG or HAAR Like).

Compare and contrast the results. You should evaluate the mAP (mean Average Precision) for Object Detection. Discuss the different properties of the video footage such as weather, video quality and time of day.

Excellent resource explaining mAP
(https://medium.com/@jonathan_hui/map-mean-average-precision-for-object-detection-45c1 21a31173)

Marks - 10

Required

Script(s) with the above specifications. Document the libraries and versions needed to run your script.

Documentation comparing the different results

Task 5 - Video Analysis

The final task involves the development of a car counting system. From the live feed develop a system that is able to count the number of cars on the road or disembarking the ship. The system should differentiate between different carriageways and detect both cars and motorcycles. Document the results and report on the accuracy of your system. Compare results of feeds for different situations such as nighttime vs daylight or weather conditions such as rainy day vs sunny.

Marks - 15

Required

Script(s) with the above specifications. Document the libraries and versions needed to run your script.

Documentation comparing the different results

Documentation

A single pdf file, properly formatted and structured, needs to be submitted to document the progress of this project. The documentation needs to have a section for each of the above tasks, clearly showing how it was implemented and evaluated/tested. Where appropriate, students are expected to properly cite third party work. The documentation should include figures and tables to provide as much information as possible about the process and support its discussion. The document must not be longer than 20 pages.

This component carries 20 marks that are allocated accordingly:

- General structure, use of English and clarity: 5 marks
- Proper acknowledgement of third party work and references: 5 marks
- Methodology and evaluation of each task: 10 marks