Video Recognition

Due 11.59 pm, Sunday 20 October 2024 (Week 11)

General description

This **INDIVIDUAL** assignment aims to consolidate your practical skills in developing image recognition systems using feature-based classifiers.

In this assignment you will **develop code** to perform object detection and tracking on a video you record and **submit a brief report** (no longer than 500 words) describing your findings and results from this task.

Important notes

1. This is an individual assignment. You must complete the tasks on your own and submit your own work only.
2. Generative AI tools such as ChatGPT **cannot** be used in this assessment task, and will be considered academic misconduct.
3. Recall that **the University takes academic misconduct very seriously**. Academic misconduct includes working with other students on an individual assignment, plagiarism, copying, and cheating. You should familiarise yourself with your responsibilities in relation to Academic Integrity and if you have any questions, direct them to your subject coordinator. Further information can be found on the website at: https://www.latrobe.edu.au/students/admin/academic-integrity.

Submission guidelines

- Submit before 11.59 pm, Sunday 20 October 2024.
- You will need to submit multiple files to the LMS:
  - Your code as a .ipynb notebook, through the assignment submission area on LMS. Obtain this file from your completed notebook on Google Colab by clicking File > Download > Download .ipynb.
  - Your recorded video as a .mp4 video file or similar.
  - Your output video showing tracking results as a .mp4 video file.
- Your code should be neat and readable. Don't leave unused code or functions in your submission. A maximum of 6 marks will be deducted for messy code with unnecessary functions or other mistakes.
- A penalty of 5% of the total available marks per day will be imposed on late submissions of assessment up to five working days. An assignment submitted more than 5 working days after the due date will not be accepted, and a zero mark will be assigned.

Tasks [40 marks]

You should prepare a brief report (no longer than 500 words) describing your findings and results from this task.

1. Record a short video approximately 15–20 seconds long using your own phone, camera, or webcam. Submit your recording as a video file named "task1.mp4" (or "task1.mkv", or any other common video format supported by OpenCV). **[6 marks]**
   - There must be at least 5 objects belonging to established MS COCO categories present in the video. Please consult the list of the 80 MS COCO class names[1].
   - The video must have some moving objects or be panning for at least a few seconds. Try to keep the motion relatively steady with smooth, slow transitions to assist with tracking.
   - Your face must be present in the video for at least a few seconds. Feel free to make a funny face if you'd like.
   - You will need to submit the video as part of the assignment, but note that the video will only be accessible to subject teaching staff for the purposes of marking the assignment.
   - Audio is not required.
2. Generate and filter detections from a pre-trained Mask R-CNN model to track the moving objects in your video. **[17 marks]**
   - Filter the detections to include only the object classes that you know are in the video. Be sure to include the "person" class.
   - Select a suitable confidence score threshold for pruning bad detections. You may have a different threshold for each object class if you find that to be beneficial.
3. Apply object tracking to the detections using either IoU or box centre distance as the association method. Create and submit a video named "task3.mp4" which displays your tracking results overlaid on your recorded video. **[17 marks]**
   - Track each object class separately. For example, it should not be possible for a single track to contain both "person" and "bird" detections.
   - Experiment with code adjustments to improve your tracking results. Make a note of all the things you try and the effects of the changes to discuss in your report.

1 MS COCO class names: https://gist.github.com/AruniRC/7b3dadd004da04c80198557db5da4bda