COMP9517 Computer Vision
Project, T1 2020

Project Synopsis

The project is worth 45% of the total course marks. Refer to the marking criteria file for detailed information about marking. Submission instructions and demo schedule will be released later.

This project consists of two components:

A. **Individual Component:** This component must be completed individually by each student. (**Individual Component is worth 15% of the total course marks**)

B. **Group Component:** This component must be completed by a team of 5 students. **(Group Component is worth 30% of the total course marks)**

Project Description

Analysing pedestrians (or humans) in videos is one of the most common applications of Computer Vision. A wide range of practical scenarios ranging from simple domestic surveillance to self-driving cars thrive on the success of pedestrian detection, tracking and advanced analysis of motion. Many well-established computer vision techniques in conjunction with machine learning techniques are useful in this task. These may include image / video pre-processing, feature extraction, classification, motion detection, tracking and recognition, and also deep-CNN algorithms.

This project aims to familiarise students with some of the major techniques involved in pedestrian detection, tracking and analysis. The tasks for each component of the project include:

1. <u>Individual Component</u>: classification of single images with and without the presence of pedestrians

2. <u>Group Component:</u> Detection of pedestrians and analysing the movements in videos

A Individual Component - 15 marks

Data: For this component you are provided with a dataset published by CSIRO [1] containing a large collection of images with (positive) and without (negative) pedestrians in them. You can download the data from the folders in ***Project_Data -> Individual_Component:***

1. Training_Data: This folder includes a subset of training data from the CSIRO database, including 3 positive sets (train_positive_A, train_positive_B and train_positive_C) and one negative set (train_negative_A)
2. Test_Data: This folder includes a subset of validation data from CSIRO database, including one positive set (test_positive) and one negative set (test_negative)
3. License.txt: License file

Task implementation: In this task you will implement a Python solution to classify positive images (with pedestrians) against negative images (without pedestrians). You can implement either a supervised or an unsupervised classification technique. If you are implementing a supervised technique, you may use either all the training data provided or a portion of it, depending on the algorithm of your choice / computational resource limitations.

Note: You are required to use traditional feature extraction techniques from computer vision (hand-crafted or engineered features, and not deep learning features) to implement this task.

Evaluation: You should evaluate your implementation with suitable metrics on the test data provided.

Deliverables for Individual Component:

Each student will submit an individual report of maximum 3 pages (2-column IEEE format), along with your source code(s) by Apr 1st, Wednesday week 7. The report should include the following parts:

1) Introduction and Background- briefly discuss your understanding of the task specification, data and a brief literature review of relevant techniques
2) Method (implementation)- justify and explain the selection of the techniques you implemented, using relevant references when necessary.
3) Experiment- explain the experimental setup and the evaluation methods and metrics used.
4) Results and Discussion- provide some visual results in addition to statistical evaluation, along with a discussion on performance and outcomes.
5) References

B Group Component - 30 marks

The group Component **consists of 3 tasks**, each of which needs to be **completed as a group and will be evaluated ONCE for the whole group**.

Data: For this component you are provided a sequence of images (extracted from a video) from The Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2009) [2] in ***Project_Data -> Group_Component***.

Task 1:

In this task you will implement a Python solution to detect and track all the pedestrian movements within the video (image sequence) frames. You will have to:

a) detect all the moving pedestrians and draw a bounding box around each of them
b) draw the trajectory (track / path) of each moving pedestrian
c) print (as an output to the terminal, or directly on the image sequence / video) the real-time count of all the moving pedestrians detected since start of the video frame 1.

Task 2:

You should allow the user to draw a rectangular bounding box within the whole window (hint: use OpenCV to draw) on the first frame / image. Then print (as output to the terminal, or print on the image sequence / video) the real-time counts of:

a) pedestrians who enter the bounding box
b) pedestrians who move out of the bounding box

Task 3:

In this task you will implement functions to enable group behaviour recognition. By analysing the image sequence (video) you should:

a) provide real-time information on how many pedestrians walk in groups (of more than 1 person), and how many pedestrians walk alone
b) detect and show (using a bounding box or other means of pointing) occurrences of group formation (eg. two people meet and stay together for more than one frame, two people meet up and walk together, etc) and group destruction (when members in a group split).

Deliverables for Group Component:
During the scheduled lecture and lab hours in week 10 (Wednesday Apr 22nd and Friday Apr 24th), each group will make a 15-minute demo/presentation to one tutor and one assessor. This should include a short slide show presentation (5 slides maximum) explaining your methods and evaluation, followed by a demonstration of your implementation. Finally, you will answer questions from the assessors/audience. The tutor and assessor will provide feedback during this session. All group members should be present for this demo. The demo schedule will be released closer to the deadline.
Note: Given the migration of all teaching related activities online, the demo/presentation is also likely to be held online, as also the question/answer session and feedback from assessors. In that case, all group members will be expected to present online and participate in the proceedings. More detailed instructions will be provided closer to the demo/presentation dates.

Each group will also submit a report of maximum 10 pages (2-column IEEE format) along with the source code(s) by Apr 24th, Friday week 10. The report should include the following parts:
1) Introduction- discuss on your understanding of the task specification and data sets
2) Literature Review- review of relevant techniques in literature, along with any necessary background to understand the techniques you selected.
3) Method (implementation)- justify and explain the selection of the techniques you implemented, using relevant references and theories when necessary.
4) Experimental setup- explain the experimental setup and evaluation methods.
5) Results and Discussion- provide statistical and visual results, along with discussion on performance and outcomes
6) Conclusion- what worked and did not work, recommendations of future work
7) Contribution of Group Members – in this section you MUST present each group member's contribution in brief. In maximum 3 lines per member, describe the component(s) that each group member contributed to.
8) References

Group Project Logistics
Due to the university's guidelines to address COVID-19 outbreak, all communication for the group project MUST be maintained on the Microsoft Teams platform. Once the groups are formed, a tutor will be assigned to each group. The tutor will create a Team in Microsoft teams for each project group, then invite group members to it. The group can then maintain in the Team all the communication, code sharing, task planning and also the consultation sessions with the assigned tutor. Some useful apps you can install in your Microsoft Teams include:
1. Github / Bitbucket for code sharing
2. Asana / Trello for task planning

More information on the usage of Teams will be made available on the course webpage in WebCMS3.

References

[1] G. Overett, L. Petersson, N. Brewer, L. Andersson and N. Pettersson, "A new pedestrian dataset for supervised learning," 2008 IEEE Intelligent Vehicles Symposium*, Eindhoven, 2008, pp. 373-378.*

[2] Ferryman, J. & Shahrokni, A. [PETS2009: Dataset and challenge](). In 11th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), 2009.

Some Useful Papers

[3] Kothiya, Shraddha V., and Kinjal B. Mistree. "A review on real time object tracking in video sequences." In *Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on*, pp. 1-4. 2015.

[4] Wang, Xiaogang. "Intelligent multi-camera video surveillance: A review." *Pattern recognition letters* 34, no. 1 (2013): 3-19.

[5] Munder, Stefan, and Dariu M. Gavrila. "An experimental study on pedestrian classification." *IEEE transactions on pattern analysis and machine intelligence* 28, no. 11 (2006): 1863-1868.

[6] Tuzel, Oncel, Fatih Porikli, and Peter Meer. "Pedestrian detection via classification on riemannian manifolds." *IEEE transactions on pattern analysis and machine intelligence* 30, no. 10 (2008): 1713-1727.

[7] Jiang, Zhengqiang, Du Q. Huynh, William Moran, Subhash Challa, and Nick Spadaccini. "Multiple pedestrian tracking using colour and motion models." In *2010 International Conference on Digital Image Computing: Techniques and Applications*, pp. 328-334. IEEE, 2010.

[8] Li, Xin, Kejun Wang, Wei Wang, and Yang Li. "A multiple object tracking method using Kalman filter." In *The 2010 IEEE international conference on information and automation*, pp. 1862-1866. IEEE, 2010.

[9] Bashir, Faisal, and Fatih Porikli. "Performance evaluation of object detection and tracking systems." In *Proceedings 9th IEEE International Workshop on PETS*, pp. 7-14. 2006.

[10] Hu, Weiming, Min Hu, Xue Zhou, Tieniu Tan, Jianguang Lou, and Steve Maybank. "Principal axis-based correspondence between multiple cameras for people tracking." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, no. 4 (2006): 663-671.

[11] Molchanov, V. V., B. V. Vishnyakov, Y. V. Vizilter, O. V. Vishnyakova, and V. A. Knyaz. "Pedestrian detection in video surveillance using fully convolutional YOLO neural network." In *Automated Visual Inspection and Machine Vision II*, vol. 10334, p. 103340Q. International Society for Optics and Photonics, 2017.