Image and Vision Computing Mini-Project Assignment

7/11/2018

In this mini-project, you will explore building a person search system for surveillance images. The system will support two features: A re-identification module, which retrieves people given a query image (used for matching the same person across disjoint camera views); and an attribute-based search system, which enables an operator to find people matching a given textual description. The assignment is broken into three sub-parts: Part 1: Person Re-identification (35% marks), Part 2: Person Attribute Recognition (35% marks), Part 3: Open-ended Bonus (30% marks).

You will do the assignment in teams, and each one has at most two members. Make your own groups and fill in the spreadsheet with your and your partner's name. If you can't find a partner you can work on your own, but you get no extra credit for doing all the work yourself. The assignment issue date is Wed. 7th Nov, and due date is 4pm, Thu. 22nd Nov. The assignment is expected to take up to 12 hours work per person in the group.

Part 1: Person Re-identification

Person re-identification (re-id) is to retrieve person images in one camera view that are the same identity as a query image in another camera view. A straightforward way to address this problem is to take the query image and a putative matching image, and define a binary classifier that reports whether the two images are the same person (+1, same), or different people (-1, different) as illustrated in (Figure 1). To train a person re-identifier system using this approach, there must be many pairs of person images, some of which are two images of the same person (+1) and some are two images of two different people (-1).

The first task of the assignment is to implement a person re-identifier. In the assignment package you are given:

- Data for this tasks are in 'person_re-id_test.mat' and 'person_re-id_train.mat' respectively (drawn from the DukeMTMC-reID dataset [2, 1])
- An example script Assignment1.m that implements a simple baseline for person re-id.

Runnining the example script, you will see that it firstly loads up the image pairs, followed by resizing each image to 128×64, extracting HoG features. In this script, a classifier (SVM) is trained to predict the similarity of any pairs of person images (same or different identity) and the accuracy of this baseline is evaluated.

By running the demo script, you will gain a basic understanding of a potential re-id system architecture. The example script obtains around 11% mean Average Precision and then your task is to improve this system (around 43.46% should be feasible) by applying the knowledge you learned in the IVC course:

- Extract better features. The HoG, LBP and SIFT suggested in the previous exercise would be a good start.
- The Color histogram feature is a good choice for person re-id.
- You can also try to enhance them by making a bag of words with normalization.

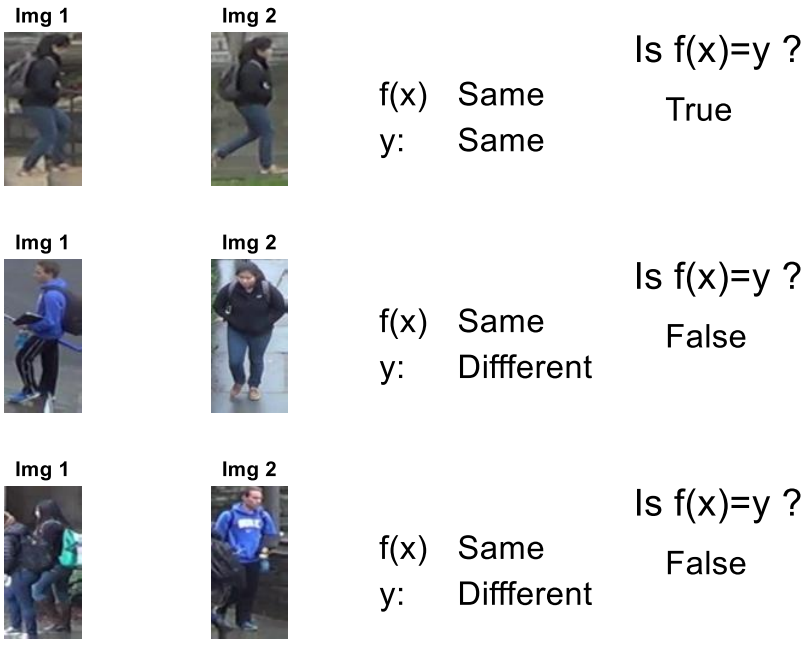

Figure 1: Some examples of person re-identification.

- Learn a better classifier. The SVM is a good starting point that is fast easy to train compared with Deep Neural Networks. You can try to tune the SVM to make it perform well (consider hyperparameter cross-validation, etc).

If you do the above and want to go further, you can try:

- Download more of the DukeMTMC-reID to obtain more pairs of training samples.
- Tune HoG, LBP, SIFT and Color histogram well
- Use deep features.

Then you should be able to get significantly better results.

Part 2: Attribute-based Person Search

Attribute-based person search addresses the problem of retrieving person images given some attributes query attributes (e.g. the gender, whether the person has backpack or bag and so on). This is often built upon attribute recognition modules (Figure 2) which are functions that inputs a person image and runs a classifier to predict the value of a categorical attribute (e.g., male or female). To train an attribute recognizer, we need many person images and the corresponding labels where the label value represents whether a person has a certain attribute (+1) or not (-1). addresses the problem of taking a person image and recognizing some attributes of the person

To achieve this, you need to implement a person attribute recognizer. In the assignment package you are given:

- Data for this task are in 'person_attribute_te.mat' and 'person_attribute_tr.mat' respectively (drawn from the DukeMTMC-reID dataset)
- An example script Assignment2.m that implements a simple baseline for person attribute recognition.

By running the example script, you will gain a basic understanding of attribute recognition. In this script, the HoG feature used together with SVM classifier to predict whether a person has a given attribute or

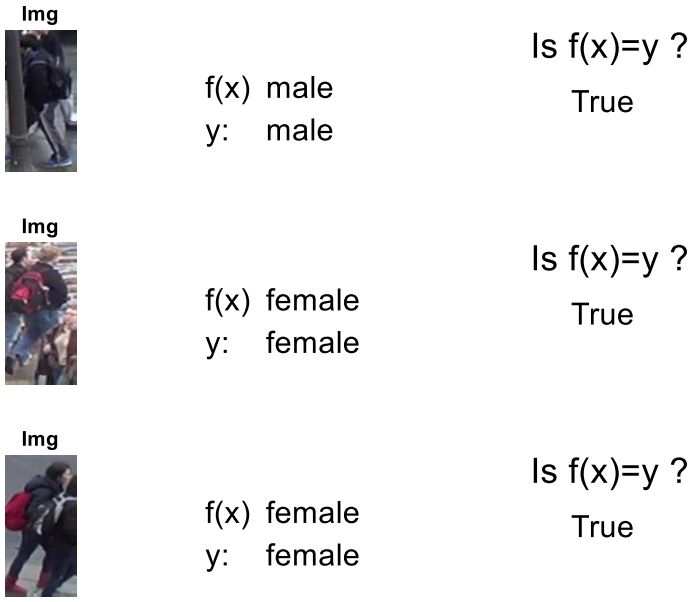

Figure 2: Some examples of attribute recognition.

not. You will see that the baseline achieves around 79% accuracy on average. However note that this is an raw accuracy number and high numbers can be achieved simply by making all negative predictions for relatively rare attributes (such as suitcase carrying). In any case our goal is to achieve an attribute-based person search capability, so our evaluation will be in terms of retrieval mAP of person images given a specific attribute query. The baseline script achieves 33% mAP. You task is to improve the baseline system provided and improve the attribute recognition and search performance (a good target to aim for is 43.49%).

To obtain a better solution you are suggested to upgrade the provided solution as follows:

- Extract better features. The HoG and LBP features suggested in the previous exercise would be a good start.
- You can also try to enhance them by making a bag of words with normalization.
- Learn a better classifier. Consider cross-validating your SVM.

If you do the above and want to go further, you can try:

- Download more of the DukeMTMC-reID to obtain more pairs of training samples.
- Tune HoG, LBP, and SVM well
- Use deep features.

Then you shoule be able to get significantly better results.

Part 3: Bonus

Once you have achieved Part 1 and 2, the final Bonus section is open-ended. Feel free to impress us! Possibilities include:

For example:

- Use deep learning to significantly outperform the target performance suggested in Parts 1 and 2.
- Put everything together into a complete person search system (Figure 3).

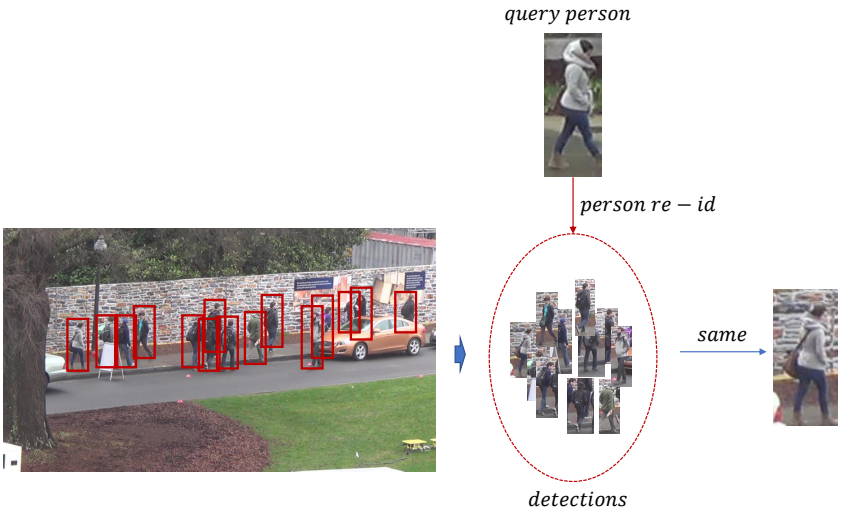

Figure 3: The pipeline of people search.

- Any other cool application of the technology. the technology.

Person Search

If you would like to try and implement a complete person search system for the bonus section, some training data and baselines are provided. Specifically, person search is to locate a query person in a raw surveillance images (contrast the Re-ID sub-problem where everything is performed on detected bounding-boxes). Given an image of a query person and a set of raw full-frame surveillance images ('gallery' images), the person search system first detects all pedestrians in all gallery images and then performs person re-identification to match the query against all the detected people. This means that both effective person detection and person matching (Re-ID) are necessary for persons search. As you have implemented a pedestrian detector in the lab exercise, the simplest way to approach this is to combine your pedestrian detector from the lab wih yor re-id system in Part 1.

In the assignment package, you are given:

- A set of cropped pedestrian images and background images are generated from the DukeMTMC-reID Dataset. This data is provided in the pedestrian_detection.mat for training pedestrian detector.
- A number of query person images and a set of gallery images are also provided in the people_search.mat file.
- A demo script Assignment3.m that shows the concept of people search is provided. You are expected to build up the people search system.

The baseline detector achieves around 1.11% mAP on pedestrian detection, which trasnlates to 0.1% mAP on people search. You can improve people search by:

- Upgrade the pedestrian detector to be a more realistic multi-scale detector.
- Build up a joint person search framework that can perform pedestrian detection and person re-identification simultaneously.
- Use deep features and download more data to train this joint framework.

You expected to illustrate how your method works on improving the performance and give some discussion on the improvement.

Submission and Grading

The grading will be based primarily on: (1) Your score on the testing set (Part 1&2). Supported by: (2) A brief report/presentation, (3) A mini demo/viva of your implementation (Parts 1-3).

Test Set: You have been provided testing set for all three parts. The assessable score will be based on the results on the testing set.

Report: The report should take the form of a short 7 page powerpoint presentation (template to be provided). Page 1: Your team members, Page 2-3: The methodology implemented for detection, the detection result obtained (test set score). Page 4-5: The methodology implemented for person re-id & attribute recognition and people search, the results obtained (testing set score). Page 6-7: Description of your attempt at bonus marks (eg advanced methodology implementation, system integration), the result or screenshot of the outcome.

Live Demo/Viva Session: Fri. 23rd Nov, (TBC). Each group will have 5 minutes to present their report slides and give a mini demo and then answer some questions of the marker/lecturer.

Submission

You should submit a zip file containing your code (source files only) and report/presentation by Thu. 22nd Nov, 4pm. The command for the DICE-based submission is:

```
submit ivc cw1 PRESENTATION CODE
```

where PRESENTATION is your (pdf or ppt) presentation, and CODE is the zip file of your code.

Good Scholarly Practice

This assignment is expected to be in your own words and code. Short quotations (with proper, explicit attribution) are allowed, but the bulk of the submission should be your own work. Use proper citation style for all citations, whether traditional paper resources or web-based materials. Please remember the University requirement as regards all assessed work for credit. Details about this can be found at:

Postgraduate-Taught and Postgraduate Research assessment regulations and feedback policy.

Academic misconduct policy.

Furthermore, you are required to take reasonable measures to protect your assessed work from unauthorised access. For example, if you put any such work on a public repository then you must set access permissions appropriately (by permitting access only to yourself, or your group in the case of group practicals).

References

[1] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara, and Carlo Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In *European Conference on Computer Vision workshop*, 2016.

[2] Zhedong Zheng, Liang Zheng, and Yi Yang. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In *International Conference on Computer Vision*, 2017.