Course work in Computer Vision
Vehicle Re-identification

What to submit: You should submit a single PDF-format report, along with the running evidence, *e.g.*, the log file. You must format your submission using the provided template and convert the final report to the PDF format for submission. This coursework is worth 25% of overall mark.

1 Introduction

Vehicle re-identification (Re-ID) is a challenging computer vision task that aims to match vehicle images captured from non-overlapping camera views. The goal is to identify the same vehicle instance across multiple camera views, which can be useful for traffic management and law enforcement applications. Vehicle Re-ID has several challenges, including variations in viewpoint, illumination, occlusion, and camera hardware differences. Researchers have proposed various approaches to tackle these challenges, including deep learning-based methods that leverage powerful feature representations to match similar vehicles. The performance of vehicle Re-ID has improved significantly in recent years, but it remains an active area of research due to its practical importance and technical difficulty.

The purpose of this coursework is to gain real-world development experience in design, training, and evaluation of a Convolution Neural Network (CNN) for vehicle Re-ID. This will be achieved through the Google Colab with PyTorch as the framework.

1.1 Required Resources for this Coursework

To do this coursework you will need to connect to Google Colab.

1.2 Getting Started

To better understand the motivation, definition, basic methods, and evaluation metrics of Re-ID, we suggest reading these two surveys before you start:

- Deep Learning for Person Re-Identification: A Survey and Outlook [3] (In particular, Sec.1 Introduction and Sec.2 Closed-World Person Re-Identification)

- Trends in Vehicle Re-Identification Past, Present, and Future: A Comprehensive Review [1] (In particular, Sec.1.3 Re-Identification, Sec.2.4 Vision-Based Vehicle Re-Identification and Sec.3 Vision-Based State-of-the-Art Vehicle Re-Identification Approaches)

We have lab sessions provided on the tutorial of PyTorch to train and test CNNs for image classification, coupled with knowledge from the associated lectures. The PyTorch Introductory lab sheet and the lecture slides from these sessions remain available on SurreyLearn in case you missed any classes. All these are helpful for this coursework assignment.

1.3 Support for Coursework Starting

We have provided a reference code base in this Github repo and a Colab demo to show how to start. For any problems, please feel free to open an issue.

2 Deliverables

You must run your project in the Colab to undertake the task in this coursework. For the report, you must document the experiments performed using your software; You should detail the method used and the numerical and qualitative results (including the running evidence).

3 Learning Outcomes

After completing this coursework along with the designated lectures, you should be able to:

1. Design and implement a convolutional neural network for vehicle Re-ID.
2. Apply appropriate training strategies to learn a discriminative CNN for representation learning.
3. Conduct a scientific investigation into fine-grained image retrieval systems in general.

4 Plagiarism

You must complete this coursework individually. If you copy code or text from the web, or another student, and include it in your project without clear attribution, then you have committed plagiarism. Undetected plagiarism degrades the quality of your degree, as it interferes with our ability to assess you and prevents you learning through properly attempting the coursework. Consequently if we suspect plagiarism you will referred to an Academic Misconduct Panel which may carry with it academic sanctions.

5 Main Assignment Tasks

During this project, we will walk you through a series of experiments in the use of CNNs for an image retrieval (e.g., vehicle re-identification) problem. Some of the main tasks are:

1. First, you need to run the demo code successfully without error and make sure fully understand every line of code.

2. You should experiment with different CNN architectures to see which gives you the highest performance. This might include the use of existing architectures (like ResNet) or modified versions of those architectures .
3. You should experiment with different hyper-parameters, *e.g.*, those parameters such as the learning rate, weight decay, optimizer and etc. you vary for model training.

6 Dataset Choice

We provide a dataset (VeRi) for this coursework [2]. It contains over 50,000 images of 776 vehicles captured by 20 cameras covering a 1.0 km^2 area in 24 hours, which makes the dataset scalable enough for vehicle Re-ID research.

6.1 Dataset Access

Cloud link: https://drive.google.com/file/d/1TfGT84sEv-KBXOwi735_ECr7gjTyCJNz/view? usp=sharing

Please don't distribute this dataset or use it for any other purposes.

---

7 Starting Early!

It takes quite time to train a CNN model. Training a ResNet will take time depending on the GPU resource. You should start this coursework early to reduce stress level for making the deadline.

8 Submission Guidance

To format your report submission, please strictly follow the provided word report template. To make the structure of report clear, you are suggested to cover the marking points one by one.

Except the report, the log file for model training and evaluation (as mentioned in each question) is needed for submission. Note, the code base we provide can generate the log file. You need to convert your final report into the PDF format for submission.

Baseline: The default settings already provided for you in the code base.

Hyperparameters: The parameters that are not learnable, and you can set before starting the training.

Important instructions:

1. Regarding log files:
    a. Please submit your own log files from your codebase.
    b. All log files are automatically watermarked and are unique to each run. Do not change the log file structure.
    c. All log files should be named based on the section number and the question number.
    (log_{Section_num}_{Question_num}.txt)

    For example, a log file generated for an experiment corresponding to question 2 in section 1, should be named as "log_1_2.txt"
2. Regarding word limit:
    a. Please ensure your answers do NOT exceed the word limit.
    b. Going beyond the word limit will be penalized.
    c. Content within Tables/graphs/log files is not counted in the word limit.
3. Regarding *presentation and clarity* of your answers:
    a. In addition to the scientific content, you will also be assessed on the presentation and clarity of the writing.
    b. This accounts for 10 marks. Following criteria to be taken into consideration:
        i. Figures should be well presented. The axis and the markings should be easily readable.
        ii. The writing should be clear, grammatically correct, and the ideas come across easily to the reader.
        iii. It is important to use tables if you're discussing results across different values of hyperparameters.

NOTE: Please read all questions carefully before attempting to answer. The below three sections account for 90 marks and (as mentioned above) 10 marks are allotted for *presentation and clarity* of the overall report.

Section 1 – Familiarity with the provided code. [40 marks]
Max. 200 words.

1. Run the code using the default settings. **[20 marks]**
    a. Provide evidence in terms of the log file.
    b. Discuss the training and evaluation process followed by implications of the observed performance using an appropriate metric. Max. 100 words.
2. Apply another CNN variant (that is not provided in the default settings). **[10 marks]**
    a. Provide evidence in terms of the log file.
    b. Critically discuss and contrast the results with what observed in question 1 above. Max. 50 words.
3. Apply one more neural network architecture (say, a transformer variant). **[10 marks]**
    a. Provide evidence in terms of the log file.
    b. Critically discuss and contrast the results with what observed questions 1 and 2 above.
    Max. 50 words.

Section 2 – Dataset preparation and Augmentation experiments. [25 marks]
Max. 250 words.

1. Apply any one data augmentation technique (for example, "crop"). Discuss the results in comparison when no data augmentation is employed, i.e. the default configuration in the provided code. **[10 marks]** Max. 100 words.
2. Apply two different augmentations in isolation (for example, only "blurring" or only "horizontal flip" etc.) and discuss the implications of each augmentation and analyze the results in comparison when no data augmentation is employed. Highlight any improvement or drop in overall score. **[10 marks]** Max. 100 words.
3. Combine augmentation techniques employed in questions 1 and 2 above (for example, "crop" + "blurring" + "vertical flip"). Highlight any improvement or drop in overall score. **[5 marks]**
    Max. 50 words.

Section 3 – Exploration of Hyperparameters [25 marks]
Max. 250 words.

1. Exploration of Learning Rate (LR). **[10 marks]**
    a. Experiment with 4 values of LR (in addition to the default value).
    b. Discuss the effects observed on overall performance.
    Max. 100 words.
2. Exploration Batch sizes. **[10 marks]**
    a. Fixing the best LR value from the experiments in question 1 above, experiment with 4 different values of the BS (in addition to the default value).
    b. Discuss the effects observed on overall performance.
    Max. 100 words.
3. Exploration of the optimizer. **[5 marks]**
    a. Fixing the best LR value and best Bath Size value from the experiments in question 1 and 2 above, respectively, experiment with changing the optimizer to SGD. (use PyTorch's internal class)
    b. Discuss the effects observed on overall performance.
    Max. 50 words.

References

[1] Jianhua Deng, Yang Hao, Muhammad Saddam Khokhar, Rajesh Kumar, Jingye Cai, Jay Kumar, and Muhammad Umar Aftab. Trends in vehicle re-identification past, present, and future: A comprehensive review. *Mathematics*, 9(24):3162, 2021. 2

[2] Xinchen Liu, Wu Liu, Huadong Ma, and Huiyuan Fu. Large-scale vehicle re-identification in urban surveillance videos. In *International Conference on Multimedia and Expo (ICME)*. IEEE, 2016. 3

[3] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven CH Hoi. Deep learning for person re-identification: A survey and outlook. *IEEE transactions on pattern analysis and machine intelligence*, 44(6):2872–2893, 2021. 2