COMP0197: Applied Deep Learning
Assessed Component 2 (Group Work – 25%)
Assessed Component 3 (Individual Report – 25%)

Submission on Moodle

Project: Weakly-supervised neural networks for segmentation

1 Introduction

Weakly-supervised segmentation is a machine learning subfield that leverages limited supervision, such as image-level labels, bounding boxes, or partial annotations, to train models that perform dense prediction tasks of semantic segmentation. This approach mitigates the high annotation cost associated with fully-supervised methods. Recent advancements include methods based on multiple instance learning, class activation maps (CAMs), self-training, and consistency regularization.

2 Project

In this project, you will implement and evaluate a weakly-supervised neural network for semantic segmentation of images (group work), then describe the methods, implementation and results in a report (individual report).

The aim of this project is to assess your ability to reason about design choices in deep learning experiments, implement deep learning frameworks, evaluate developed models, work in a group and summarise and communicate your work in a technical report.

2.1 The data set

The Oxford-IIIT Pet Data-set will be used for training and evaluation in this project. The dataset contains thousands of images with pixel-level annotations for animal segmentation. For weakly-supervised learning, only image-level labels and/or bounding box annotations will be used as primary supervision during training. Additional unlabelled data or pseudo-labelling methods may be considered to improve performance.

2.2 The "minimum required project"

The so-called minimum required project (MRP) (70%) for this group work consists of the following points, which are expected to be found in your submission. Some of these points are discussed in more details to get you started.

- Formulate a specific weak supervision problem to address and justify its usefulness and feasibility.
- Discuss and justify which weakly-supervised segmentation algorithm to use.
    - Use references and citations to support your choice.
- Optionally, motivate, identify and collect additional weakly-labelled or unlabelled data.
- Implement a weakly-supervised segmentation framework.
- Design and conduct experiments for network comparisons, at least:
    - Compare the framework with baseline models trained using fully-supervised methods.

  - Compare the benefits of the weak supervision, between different configurations of the framework (i.e. an ablation study by varying important hyperparameters).
- Describing the implemented methods and conducted experiments.
- Summarising obtained results and drawing conclusions.

2.3 An "open-ended question"

To be awarded the remaining 30% in this project, you need to propose a novel study question (an open-ended question, OEQ). Novelty is encouraged in this part. The group should clearly identify and generate a hypothesis, design experiments, and analyse the obtained results for a quantitative conclusion. This part should build on the MRP and be relevant to weakly-supervised segmentation algorithms. Describe the question, experiment and results clearly and cohesively with the MRP in the report.

Some example study questions are given as follows.

- *How does the quality of weak annotations affect segmentation performance?*
- *Can any additional self-supervision or unsupervised approaches improve the performance of weakly-supervised segmentation?*
- *How do different forms of weak supervision (e.g., bounding boxes vs. image-level labels) impact results?*
- *What has been learned from weak supervision compared to full supervision?*

3 Submission and Marking Criteria

Each group will submit one bundle of code in a single zip file and an accompanying instruction PDF file (50 MB file size limit applies). Individual will submit a report in PDF format. Both should be submitted before the same deadline.

3.1 Code

In this project, you must only use one of the PyTorch or TensorFlow for the entirety of the project. However, you can use up to **three additional pip-installable packages** within the "comp0197-cw1-xx" conda environment, in addition to the available tools in the conda environment.

A "instruction.pdf" file is required to

1. List additional installed packages (if any) within the "comp0197-cw1-xx" environment.
2. Provide detailed steps to run the code and reproduce all reported results.

It is your responsibility to ensure the compatibility of the packages you added and that they can run on a marking environment on Ubuntu 22.04. There is no specific format or style required for how you design, structure and implement your code. However, general good programming practice and readability of the code may contribute towards meeting the requirements in marking criteria below.

Generative AI (GenAI) tools can also be used in an assistive role in this coursework, in accordance with the UCL regulations. If used, a statement must be included in each submitted Python file detailing their role in the coursework. Further UCL guidance can also be found from the above link. It is your responsibility to ensure the code runs following the provided instruction.

3.2 Report

Each group member must submit one individual report. Unlike the first individual coursework, the marking is primarily based on the report. The submitted report must contain the following sections (followed with the suggested content), structured as a research paper. The percentages indicate the maximum marks for each section, each with a 7:3 split for the MRP and OEQ, respectively.

Note: your report must not exceed 6 pages in the LNCS template, excluding references. Attempts to gain extra space or pages, for example by modifying fonts or margins, will be considered as a case of exceeding word/page limit, which will be penalised according to the UCL regulations.

- Title: an interesting and informative title of the study (0%)
- Introduction: background, literature, motivation, the MRP and OEQ study questions to address (20%)
- Methods: the methodological details for the implemented weakly-supervised algorithms and other networks, optional novel methodology for the OEQ (25%)
- Experiments: MRP comparison and ablation experiments, and other experiments for the OEQ (25%)
- Results: a summary of results, comparison, ablation, quantitative analysis (20%)
- Discussion: key findings, unanswered questions, limitations and future directions (5%)
- Conclusion: summarising the study and the results (5%)

3.3 Marking

The marking criteria used is adapted using the departmental project marking criteria (both UG and PG) as reference. To summarise, in descending order of importance:

- Scientific soundness: reasoning and justification of problems, methods and experiments, background reading and understanding, experiment design and quality.
- Technical accuracy: appropriate and correct use of terminology, methodology, data, code and other tools.
- Completeness: objective achieved, conclusion of study question, completeness of report.
- Presentation: writing organisation, clarity, illustrative report, structure, code readability.
- Critical appraisal: conclusive results, informative analysis, future outlook.
- Novelty (only relevant to OEQ): this point can be subjective, and can be particularly related to the background and existing literature.

3.4 Peer assessment

There is no mandatory peer assessment. The same grade on the code from each group will be assigned to all group members, while scores on reports are individually marked. However, cases of negligence or any other academic malpractice, if reported, will be investigated and assessed per UCL regulations, on a case-by-case basis.